

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:09:56 ; Search time 85.5 Seconds
(without alignments)
83.913 Million cell updates/sec

Title: US-10-066-965A-1
Perfect score: 129
Sequence: 1 QVSLWALGWRWLRRYGNM 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1980s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	20	2 AAR89346	Aar89346 Cdk2-inte
2	129	100.0	20	2 AAW32121	Aaw32121 Interacti
3	129	100.0	20	3 AAY49340	Aay49340 Cdk2 inte
4	117	90.7	20	5 AAO18008	Aao18008 Intracell
5	97.5	75.6	17	5 AAO18007	Aao18007 Intracell
6	65	50.4	118	4 ABG24883	Abg24883 Novel hum
7	60.5	46.9	78	8 ADG22339	Adg22339 Cyanophag
8	60	46.5	2115	6 ABP70920	Abp70920 Arabidops
9	60	46.5	2150	6 ABP70925	Abp70925 Arabidops
10	60	46.5	2159	6 ABP70924	Abp70924 Maize DEK
11	59	45.7	73	4 AAU17861	Aau17861 Novel hum
12	59	45.7	73	7 ADG41241	Adg41241 Human res
13	56	43.4	122	8 ADG22291	Adg22291 Cyanophag
14	56	43.4	1086	7 ABP76305	Abp76305 Pseudomon
15	55	42.6	599	4 ABG18132	Abg18132 Novel hum
16	53	41.1	258	4 AAU36364	Aau36364 Pseudomon
17	53	41.1	258	6 ABU38632	Abu38632 Protein e
18	53	41.1	260	7 ABG74240	Abg74240 Pseudomon
19	53	41.1	382	7 ABG083419	Abg083419 Pseudomon
20	52.5	40.7	1031	8 ADM72293	Adm72293 Equine TL
21	52	40.3	62	4 AAU45246	Aau45246 Propionib
22	52	40.3	62	6 ABM41765	Abm41765 Propionib
23	52	40.3	105	8 ABG058643	Abg058643 Human gen
24	52	40.3	358	5 ABG90371	Abg90371 Human pol
25	52	40.3	431	5 ABP59544	Abp59544 Human his

ALIGNMENTS

RESULT 1
AAR89346
ID AAR89346 standard; peptide; 20 AA.
XX AAR89346;
AC
XX
DT 10-SEP-1996 (first entry)
XX
DE Cdk2-interacting peptide isolated using interaction trap assay.
XX
KW Cdk2; cyclin dependent kinase 2; assay; identifying; isolating;
cell cycle; interaction; antagonist; conformationally-constrained;
agonist; interaction trap; thioredoxin; LexA; two-hybrid system.
XX
OS Synthetic.
XX
FN WO9602561-A1.
XX
PD 01-FEB-1996.
XX
PF 20-JUL-1995; 95WO-US009307.
XX
PR 20-JUL-1994; 94US-00278082.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX (GEWY) GENETICS INST INC.
XX
PI Brent R, McCooy JM, Jessen TH, Xu C;
XX
DR WPI; 1996-105852/11.
XX
PT Interaction trap systems using conformationally-constrained proteins -
PT useful for detection of protein interactions and for identification and
PT isolation of interacting proteins.
XX
PS Claim 66; Page 60; 73pp; English.
XX
CC AAR89340-R89351 are Cdk-2 interacting peptides identified by an
CC interaction trap assay using conformationally-constrained proteins. The
CC assay comprises providing a host cell (esp. a yeast cell) which contains:
CC (i) a reporter gene operably linked to a DNA-binding protein recognition
CC site, pref. LexA; (ii) a 1st fusion gene expressing Cdk-2 covalently
CC bonded to a LexA binding protein; and (iii) a 2nd fusion gene expressing
CC a conformationally-constrained (pref. with thioredoxin) protein (either
CC expected to interact with Cdk-2 or a random peptide) covalently bonded to
CC a gene-activating moiety, and measuring expression of the reporter gene
CC as a measure of interaction between Cdk2 and the conformationally
CC constrained protein. The same system may be applied to find proteins that

26 52 40.3 1725 3 AAB48460
27 51 39.5 60 4 AAU57671
28 51 39.5 60 6 ABM54190
29 51 39.5 68 4 AAM80357
30 51 39.5 78 5 ABP01028
31 51 39.5 89 4 ABG06199
32 51 39.5 105 4 ABG11452
33 51 39.5 261 6 ABU22728
34 51 39.5 480 7 ABO82724
35 50.5 39.1 136 6 ABR56763
36 50.5 39.1 300 6 ABU21414
37 50.5 39.1 336 7 ABO82339
38 50.5 39.1 513 8 ADP98943
39 50.5 39.1 538 6 ABU21411
40 50 38.8 36 4 AAM87561
41 50 38.8 41 7 ADC78292
42 50 38.8 266 6 ADA34903
43 50 38.8 743 3 AAB11535
44 49.5 38.4 391 3 AAG44312
45 49.5 38.4 391 5 ABB93906

Aab48460 Rat lamin
Aau57671 Propionib
Abm54190 Propionib
Aam80357 Human hae
Abp01028 Human ORF
Abg06199 Novel hum
Abg11452 Novel hum
Abu22728 Protein e
Abo82724 Pseudomon
Abr56763 Human sec
Abu21414 Protein e
Abo82339 Pseudomon
Adp98943 C. albica
Aam87561 Human imm
Aac78292 Human mig
Ada34903 Acinetoba
Aab11535 SEN virus
Aag44312 Arabidops
Abb93906 Herbicida

CC interact with any other protein of interest (e.g. Ras) or proteins having
CC agonist or antagonist activity on such interactions
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 129; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVMSLWALGWRWLRYYGNM 20
Db 1 QVMSLWALGWRWLRYYGNM 20

RESULT 2

AAW32121
ID AAW32121 standard; peptide; 20 AA.

AC AAW32121;

XX 22-APR-1998 (first entry)

DT Interaction trap system aptamer 7.

DE Protein interaction; interaction trap; fusion protein; mimetic;
KW therapeutic; detection; reporter gene.

XX Synthetic.

XX WO9738127-A1.

PN 16-OCT-1997.

PD 09-APR-1997; 97WO-US005793.

XX 09-APR-1996; 96US-00630052.

XX (GEHO) GEN HOSPITAL CORP.

PA (GEMY) GENETICS INST INC.

XX Brent R, McCoy JM, Jessen TH, Xu CW;

PI WPI; 1997-512733/47.

XX New trap system for detecting protein interactions - comprises a reporter
PT gene linked to a DNA-binding-protein recognition site and fusion proteins
PT to test for interactions.

PS Claim 73; Page 52; 89pp; English.

XX Aptamers AAW32116-W32132 have been isolated from a peptide library and
CC are used in a novel interaction trap method for detecting protein
CC interactions and isolating novel proteins. The method involves a host
CC cell containing a reporter gene operably linked to a DNA-binding-protein
CC (DBP) recognition site, a fusion gene capable of expressing a fusion
CC protein which is able to specifically bind to the DBP recognition site
CC and a second fusion gene which expresses a second fusion protein which is
CC conformationally constrained and bonded to a gene activating moiety.
CC Measuring expression of the reporter gene gives a measure of the
CC interaction between p1 and p2. This method can be used to identify
CC agonists or antagonists for use as therapeutic molecules or for the
CC design of simple molecule mimetics. The method is specifically used to
CC detect an interacting protein in a population of proteins or to identify
CC a candidate interactor. Using conformationally constrained proteins can
CC provide for tertiary structural analysis and can also protect proteins
CC from cellular degradation and/or increase the protein's solubility,
CC and/or otherwise alter the capacity of the candidate interactor to
CC interact

XX Sequence 20 AA;

Query Match 100.0%; Score 129; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVMSLWALGWRWLRYYGNM 20
Db 1 QVMSLWALGWRWLRYYGNM 20

RESULT 3

AAAY49340
ID AAY49340 standard; peptide; 20 AA.

XX AC AAY49340;

XX 14-MAR-2000 (first entry)

DE Cdk2 interacting peptide i5-4.

XX Saccharomyces; mammalian; fusion protein; interactor peptide;
KW conformation-constraining protein; DNA binding moiety; Cdk2;
XX gene activating moiety; protein interaction; gene purification.

OS Synthetic.

XX US6004746-A.

PN 21-DEC-1999.

XX 20-JUL-1995; 95US-00504538.

XX 20-JUL-1994; 94US-00278082.

XX (GEHO) GEN HOSPITAL CORP.

XX (GEMY) GENETICS INST INC.

XX McCoy JM, Jessen TH, Brent R;

XX WPI; 2000-072059/06.

XX Population of Saccharomyces and/or mammalian cells comprising recombinant
PT DNA encoding fusion proteins, useful for detecting protein interactions.

PS Disclosure; Fig 3B; 24pp; English.

XX The invention relates to a population of Saccharomyces and/or mammalian
CC cells comprising recombinant DNA molecules encoding fusion proteins, each
CC consisting of a candidate interactor peptide, a conformation-constraining
CC protein and a DNA binding moiety and/or gene activating moiety. The cells
CC are useful for detecting protein interactions. The cells may also be used
CC in a method for identifying and purifying genes encoding a wide range of
CC useful proteins based on their physical interaction with a second
CC polypeptide

XX Sequence 20 AA;

Query Match 100.0%; Score 129; DB 3; Length 20;

Best Local Similarity 100.0%; Pred. No. 7.1e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVMSLWALGWRWLRYYGNM 20
Db 1 QVMSLWALGWRWLRYYGNM 20

RESULT 4

AAO18008

ID AAO18008 standard; peptide; 20 AA.

XX AC AAO18008;

XX 30-AUG-2002 (first entry)

DE Intracellular target molecule property modulation method aptamer 10M.

KW	Intracellular target; cellular component; property modulation;
KW	antimicrobial; immunomodulatory; nootropic; neuroprotective; metabolic;
KW	neuroleptic; cytostatic; cardiant; infection; immunological disorder;
KW	neurological disorder; metabolic disorder; psychiatric disorder;
XX	myopathy; cancer; cardiovascular disorder.
OS	Unidentified.
EP1205191-A1.	
XX	
XX	15-MAY-2002.
XX	
PP	13-NOV-2000; 2000EP-00403156.
XX	
PP	13-NOV-2000; 2000EP-00403156.
XX	
PPA	(CNRS) CENT NAT RECH SCI.
PPA	(MASS-) MASSACHUSETTS GEN HOSPITAL.
PPA	(MOLE-) MOLECULAR SCI INST.
XX	
PI	Colas P, Brent R, Cohen BA;
XX	
XX	WPI; 2002-418829/45.
XX	
PPT	Process for specifically modulating the properties of an intracellular
PT	target molecule used for the treatment of various disorders.
XX	
PS	Example 1; Fig 1; 33pp; English.
XX	
CC	The present invention relates to a process for specifically modulating
CC	the properties of an intracellular target molecule T, and/or of a
CC	cellular component C which interacts directly or indirectly in a cell
CC	with the target. The process involves the introduction into the cell of a
CC	chimeric molecule known as a targeted effector, comprising a recognition
CC	moiety capable of recognising T and an effector moiety. The chimeric
CC	protein or nucleic acid can be used in the preparation of a medicament
CC	for the treatment of microbial infections, immunological disorders,
CC	neurological disorders, metabolic disorders, psychiatric disorders,
CC	myopathies, genetic disorders, cancer, cardiovascular disorders and
CC	dental disorders. The present sequence is a mutant of a known anti-Cdk2
CC	aptamer used in the exemplification of the invention
XX	
XX	Sequence 20 AA;
SQ	
Query Match	90.7%; Score 117; DB 5; Length 20;
Best Local Similarity	90.0%; Pred. No. 2.4e-08;
Matches	18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	1 QVWSLWALGWRWLRRYGWNM 20
Db	1 QVWSLWALGWRWLRRYGWNM 20
RESULT 5	
AAO18007	
ID	AAO18007 standard; peptide; 17 AA.
XX	
AC	AAO18007;
XX	
DT	30-AUG-2002 (first entry)
XX	
DE	Intracellular target molecule property modulation method aptamer 10.
XX	
KW	Intracellular target; cellular component; property modulation;
KW	antimicrobial; immunomodulatory; nootropic; neuroprotective; metabolic;
KW	neuroleptic; cytostatic; cardiant; infection; immunological disorder;
KW	neurological disorder; metabolic disorder; psychiatric disorder;
KW	myopathy; cancer; cardiovascular disorder.
XX	
OS	Unidentified.
XX	
XX	EP1205191-A1.
XX	

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 55242; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 118 AA;
SQ
Query Match 50.4%; Score 65; DB 4; Length 118;
Best Local Similarity 66.7%; Pred. No. 0.65;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 3 WSLWALGWRLLRRYG 17
Db 52 WALWSRGWLLRLRG 66
RESULT 7
ADG22339
ID ADG22339 standard; protein; 78 AA.
XX
XX ADG22339;
XX
XX 26-FEB-2004 (first entry)
XX
XX Cyanophage S-2L encoded protein #84.
XX genome; cyanophage; 2; 6-diaminopurine; chemotherapy; AIDS.
XX
XX Cyanophage S-2L.
XX
XX FR2839079-A1.
XX
XX 31-OCT-2003.
XX
XX 30-APR-2002; 2002FR-00005424.
XX
XX 30-APR-2002; 2002FR-00005424.
XX
XX (INSP) INST PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX (GENO-) GENOSCOPE CENT NAT SEQUENCEAGE GRP INTERE.
XX
XX Marliere P, Kaminski PA, Galisson F, Bourzon M, Pochet S;
XX Weissenbach J, Saurin W, Robert C, Vico V;
XX
XX WPI; 2004-045746/05.
XX N-PSDB; ADG22255.
XX

PT New genomic sequence for cyanophage S-2L, useful for identifying genes
PT for synthesis of 2,6-diaminopurine bases or polynucleotides containing
PT them.
XX
XX Claim 6; SEQ ID NO 85; 423pp; French.
XX
XX The invention relates to the entire genome of cyanophage S-2L, and to the
CC protein encoded by it. Genes isolated from the genome of S-2L are useful
CC for preparing enzymes for synthesis of D-bases (D = 2,6-diaminopurine),
CC particularly D, dUMP and dTTP, or polynucleotides containing these bases,
CC polymerases involved in metabolism of D-bases and deoxynucleotide
CC analogs, for chemotherapy of AIDS. The genes, and encoded polypeptides,
CC can be used for detection and/or identification of S-2L, and for
CC identifying agents that modulate synthesis of D-bases or polynucleotides
CC containing them, and fusions of S-2L polypeptides with an antigen can be
CC used to raise specific antibodies, useful for detecting S-2L. This
CC sequence corresponds to one of the proteins encoded by the cyanophage S-
CC 2L genome.
XX
XX Sequence 78 AA;
SQ
Query Match 46.9%; Score 60.5; DB 8; Length 78;
Best Local Similarity 60.0%; Pred. No. 1.6;
Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;
QY 3 WSLWALGWRLLRRYG 17
Db 44 WALWS-AWRWLRWSWG 57
RESULT 8
ABP70920
ID ABP70920 standard; protein; 2115 AA.
XX
XX ABP70920;
XX
XX 26-AUG-2003 (first entry)
XX
XX Arabidopsis calpain.
XX
XX Plant; seed; grain; aleurone cell; transgenic plant; gene therapy;
XX calpain.
XX
XX Arabidopsis thaliana.
XX
XX WO2003011015-A2.
XX
XX 13-FEB-2003.
XX
XX 30-JUL-2002; 2002WO-US024438.
XX
XX 02-AUG-2001; 2001US-0309719P.
XX 25-OCT-2001; 2001US-0337444P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Tarczynski MC, Olsen O, Shen B, Lid SE, Li C, Jung R, Gruis DB;
XX Ananiev E, Nichols SE, Lorentzen JA, Wang CW;
XX
XX WPI; 2003-248094/24.
XX N-PSDB; ACC42644.
XX
XX Altering seed characteristics for producing transgenic plants by
PT introducing into a plant cell a recombinant expression cassette having a
PT polynucleotide that alters the number, type and configuration of aleurone
PT cells within seed.
XX
XX Disclosure; Page 75-80; 171pp; English.
XX
XX The present invention relates to a method for altering seed
CC characteristics. The method comprises introducing (into a plant cell) a
CC recombinant expression cassette comprising a polynucleotide whose
CC expression, alone and in combination with additional polynucleotides,
CC

CC alters the number, type and configuration of aleurone cells within seed,
 CC (the polynucleotide is operably linked to a promoter). The method is
 CC useful for altering seed characteristics to produce a transgenic having
 CC seeds and grains with improved nutritional, industrial and agricultural
 CC traits. The present sequence is the protein sequence for calpain, which
 CC was used to illustrate the invention

XX SQ Sequence 2115 AA;

Query Match 46.5%; Score 60; DB 6; Length 2115;
 Best Local Similarity 56.2%; Pred. No. 51;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 WSLWALGWRWLRRYGW 18
 DB 28 WILWAVNRPWRLYSW 43

RESULT 9
 ABP70925
 ID ABP70925 standard; protein; 2150 AA.

XX AC
 XX ABP70925;
 XX 26-AUG-2003 (first entry)

XX DE Arabidopsis DEK1.

XX plant; seed; grain; aleurone cell; transgenic plant; gene therapy; dek1;
 KW aleurone cell differentiation.

XX OS Arabidopsis thaliana.

XX PN WO2003011015-A2.

XX PD 13-FEB-2003.

XX PF 30-JUL-2002; 2002WO-US024438.

XX PR 02-AUG-2001; 2001US-0309719P.

XX PR 25-OCT-2001; 2001US-0337444P.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Tarczynski MC, Olsen O, Shen B, Lid SE, Li C, Jung R, Gruis DB;
 PI Ananiev E, Nichols SE, Lorentzen JA, Wang CW;

XX WPI: 2003-248094/24.
 XX N-PSDB; ACC42661.

XX Altering seed characteristics for producing transgenic plants by
 PT introducing into a plant cell a recombinant expression cassette having a
 PT polynucleotide that alters the number, type and configuration of aleurone
 PT cells within seed.

XX Disclosure; Page 149-154; 171pp; English.

XX The present invention relates to a method for altering seed
 CC characteristics. The method comprises introducing (into a plant cell) a
 CC recombinant expression cassette comprising a polynucleotide whose
 CC expression, alone and in combination with additional polynucleotides,
 CC alters the number, type and configuration of aleurone cells within seed,
 CC (the polynucleotide is operably linked to a promoter). The method is
 CC useful for altering seed characteristics to produce a transgenic having
 CC seeds and grains with improved nutritional, industrial and agricultural
 CC traits. The present sequence is Arabidopsis thaliana DEK1, which was used
 CC to illustrate the invention. The Dek1 protein is involved in aleurone
 CC cell differentiation

XX SQ Sequence 2150 AA;

Query Match 46.5%; Score 60; DB 6; Length 2150;
 Best Local Similarity 56.2%; Pred. No. 52;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 3 WSLWALGWRWLRRYGW 18
 DB 28 WILWAVNRPWRLYSW 43

RESULT 10

ABP70924
 ID ABP70924 standard; protein; 2159 AA.

XX AC ABP70924;

XX 26-AUG-2003 (first entry)

XX DE Maize DEK1 from B73.

XX plant; maize; seed; grain; aleurone cell; transgenic plant; gene therapy;
 KW dek1; aleurone cell differentiation.

XX OS Zea mays.

XX FH Key Location/Qualifiers
 XX FT Domain 1708..2013
 XX FT /label= Cystein_proteinase_domain_II

XX PN WO2003011015-A2.

XX PD 13-FEB-2003.

XX PF 30-JUL-2002; 2002WO-US024438.

XX PR 02-AUG-2001; 2001US-0309719P.

XX PR 25-OCT-2001; 2001US-0337444P.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Tarczynski MC, Olsen O, Shen B, Lid SE, Li C, Jung R, Gruis DB;
 PI Ananiev E, Nichols SE, Lorentzen JA, Wang CW;

XX WPI: 2003-248094/24.
 XX N-PSDB; ACC42659.

XX Altering seed characteristics for producing transgenic plants by
 PT introducing into a plant cell a recombinant expression cassette having a
 PT polynucleotide that alters the number, type and configuration of aleurone
 PT cells within seed.

XX Example 9; Page 105-109; 171pp; English.

XX The present invention relates to a method for altering seed
 CC characteristics. The method comprises introducing (into a plant cell) a
 CC recombinant expression cassette comprising a polynucleotide whose
 CC expression, alone and in combination with additional polynucleotides,
 CC alters the number, type and configuration of aleurone cells within seed,
 CC (the polynucleotide is operably linked to a promoter). The method is
 CC useful for altering seed characteristics to produce a transgenic having
 CC seeds and grains with improved nutritional, industrial and agricultural
 CC traits. The present sequence is maize DEK1 from B73, which was used to
 CC illustrate the invention. The DEK1 protein is involved in aleurone cell
 CC differentiation

XX SQ Sequence 2159 AA;

Query Match 46.5%; Score 60; DB 6; Length 2159;
 Best Local Similarity 56.2%; Pred. No. 52;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 WSLWALGWRWLRRYGW 18
 DB 29 WILWAVNRPWRLYSW 44

RESULT 11
AAU17861
ID AAU17861 standard; protein; 73 AA.
XX
AC AAU17861;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel human respiratory antigen #177.
XX
KW Human; respiratory antigen; respiratory disorder; throat disorder;
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW respiratory active.
XX
OS Homo sapiens.
XX
PN WO200155448-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001333.
XX
31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216980P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.

PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 28-SEP-2000; 2000US-0235935P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239537P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244647P.
PR 08-NOV-2000; 2000US-0244675P.
PR 08-NOV-2000; 2000US-0244676P.
PR 08-NOV-2000; 2000US-0244677P.
PR 08-NOV-2000; 2000US-0244678P.
PR 08-NOV-2000; 2000US-0245233P.
PR 08-NOV-2000; 2000US-024524P.
PR 08-NOV-2000; 2000US-024525P.
PR 08-NOV-2000; 2000US-024526P.
PR 08-NOV-2000; 2000US-024527P.
PR 08-NOV-2000; 2000US-024528P.
PR 08-NOV-2000; 2000US-024532P.
PR 08-NOV-2000; 2000US-0245609P.
PR 08-NOV-2000; 2000US-0245610P.
PR 08-NOV-2000; 2000US-0245611P.
PR 08-NOV-2000; 2000US-0245613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-00764860.
PR 17-JAN-2001; 2001US-00764860.
PR 14-FEB-2002; 2002US-00074095.

XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-902033/82.
XX N-PSDB; ADG40949.
XX Novel respiratory system antigen and polynucleotides encoding the
XX polypeptides, useful for treating diagnosing, treating or preventing
XX tonsillitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
XX cancer.
XX Claim 11; SEQ ID NO 479; 236pp; English.
XX The invention describes an isolated polypeptide (I) comprising an amino
XX acid sequence that is at least 90% identical to polypeptide fragment of
XX any one of 299 respiratory system antigen sequences (PS) and having
XX biological activity, polypeptide domain or epitope of PS, full-length
XX protein of PS, or variant, allelic variant or species homolog of PS. (I)
XX or a polynucleotide (II) encoding (I) is also useful for diagnosing a
XX pathological condition or a susceptibility to a pathological condition in
XX a subject which involves determining the presence or absence of mutation
XX in (II) or determining the presence or amount of expression of (I) in a
XX biological sample and diagnosing a pathological condition based on the
XX result. The human respiratory system associated polynucleotides, the
XX polypeptides encoded by them, and antibodies that immunospecifically bind
XX these polypeptides are useful in diagnosis, treatment, prevention and/or
XX prognosis of disorders of respiratory system such as throat disorders
XX (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
XX (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic
XX pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,
XX sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
XX cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
XX cancer of the nose). The polynucleotides are useful in gene therapy

Query Match 45.7%; Score 59; DB 7; Length 73;
Best Local Similarity 47.1%; Pred. No. 2.4;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 2 VWSLWALQWRWLRYYGW 18
Db 16 IWSMWPGEQWLRFPSW 32

RESULT 13
ADG22291
ID ADG22291 standard; protein; 122 AA.
XX AC ADG22291;
XX 26-FEB-2004 (first entry)
XX Cyanophage S-2L encoded protein #36.
XX genome; cyanophage; 2; 6-diaminopurine; chemotherapy; AIDS.
XX Cyanophage S-2L.
XX FR2839079-Al.
XX 31-OCT-2003.
XX 30-APR-2002; 2002FR-00005424.
XX 30-APR-2002; 2002FR-00005424.
XX (INSP) INST PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX (GENO-) GENOSCOPE CENT NAT SEQUENCAGE GRP INTERE.
XX Marliere P, Kaminski PA, Galisson F, Bouzon M, Pochet S;
XX Weissenbach J, Saurin W, Robert C, Vico V;

XX WPI: 2004-045746/05.
 DR N-PSDB; ADG22255.
 XX
 PT New genomic sequence for cyanophage S-2L, useful for identifying genes
 PT for synthesis of 2,6-diaminopurine bases or polynucleotides containing
 PT them.
 XX
 PS Claim 6; SEQ ID NO 37; 423pp; French.
 XX
 CC The invention relates to the entire genome of cyanophage S-2L, and to the
 CC protein encoded by it. Genes isolated from the genome of S-2L are useful
 CC for preparing enzymes for synthesis of D-bases (D = 2,6-diaminopurine),
 CC particularly D, dUMP and dTMP, or polynucleotides containing these bases,
 CC polymerases involved in metabolism of D-bases and deoxynucleotide
 CC analogs, for chemotherapy of AIDS. The genes, and encoded polypeptides,
 CC can be used for detection and/or identification of S-2L, and for
 CC identifying agents that modulate synthesis of D-bases or polynucleotides
 CC containing them, and fusions of S-2L polypeptides with an antigen can be
 CC used to raise specific antibodies, useful for detecting S-2L. This
 CC sequence corresponds to one of the proteins encoded by the cyanophage S-
 CC 2L genome.
 XX
 SQ Sequence 122 AA;
 Query Match 43.4%; Score 56; DB 8; Length 122;
 Best Local Similarity 80.0%; Pred. No. 9.6;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 9 GMRWLRRYGW 18
 DB 112 GMRWCRGGW 121
 |||||
 |||||
 RESULT 14
 ABO76305
 ID ABO76305 standard; protein; 1086 AA.
 XX
 AC ABO76305;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #8480.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US6551795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX WPI; 2003-615309/58.
 DR N-PSDB; ABD09876.
 XX
 CC Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 25051; 455pp; English.
 XX
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,

CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 1086 AA;
 Query Match 43.4%; Score 56; DB 7; Length 1086;
 Best Local Similarity 62.5%; Pred. No. 85;
 Matches 10; Conservative 0; Mismatches 2; Indels 4; Gaps 1;
 QY 3 MSILWLGWELRRYGW 18
 DB 52 WSRWA----WLRNGW 63
 |||||
 |||||
 RESULT 15
 ABG18132
 ID ABG18132 standard; protein; 599 AA.
 XX
 AC ABG18132;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #18123.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS82319.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 48491; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG0377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 599 AA;

Query Match 42.6%; Score 55; DB 4; Length 599;
 Best Local Similarity 60.0%; Pred. No. 63;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 4 SLWALGWRWLRYYGW 18
 | : ||| |||
 Db 171 SVERLGWYLSRVGW 185

Search completed: October 26, 2004, 06:45:08
 Job time : 86.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:20:00 ; Search time 22 seconds
(without alignments)
60.289 Million cell updates/sec

Title: US-10-066-965A-1

Perfect score: 129

Sequence: 1 QVSLWALGWRWLRYYGNM 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66319000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	20	3	US-08-504-538A-12
2	129	100.0	20	3	US-08-630-052-12
3	129	100.0	20	5	PCT-US95-09307-12
4	56	43.4	1086	4	US-09-252-991A-25051
5	53	41.1	260	4	US-09-252-991A-22986
6	53	41.1	382	4	US-09-252-991A-32165
7	53	41.1	550	2	US-08-816-155B-44
8	53	41.1	550	3	US-09-079-587-44
9	52	40.3	1725	4	US-09-560-385A-10
10	51	39.5	480	4	US-09-252-991A-31470
11	50.5	39.1	396	4	US-09-252-991A-31085
12	50	38.8	266	4	US-09-328-352-6190
13	49	38.0	61	2	US-08-637-759B-298
14	49	38.0	61	3	US-08-871-355A-298
15	49	38.0	61	3	US-09-201-945-298
16	49	38.0	135	4	US-09-270-767-31648
17	49	38.0	135	4	US-09-270-767-46865
18	49	38.0	282	4	US-09-247-890-16
19	49	38.0	282	4	US-09-724-969-16
20	49	38.0	282	4	US-09-724-852-16
21	49	38.0	318	2	US-08-619-362A-9
22	48.5	37.6	37	4	US-09-270-767-36409
23	48.5	37.6	37	4	US-09-270-767-51626
24	48	37.2	266	4	US-09-252-991A-15128
25	47.5	36.8	109	4	US-09-248-796A-25219
26	47.5	36.8	115	4	US-09-489-039A-9146
27	47.5	36.8	307	4	US-09-489-039A-12110

28	47	36.4	182	4	US-09-252-991A-17896	Sequence 17896, A
29	47	36.4	277	4	US-09-252-991A-17567	Sequence 17567, A
30	47	36.4	471	4	US-09-489-039A-7878	Sequence 7878, Ap
31	47	36.4	474	4	US-09-252-991A-26967	Sequence 26967, A
32	46.5	36.0	187	4	US-09-252-991A-26696	Sequence 26696, A
33	46.5	36.0	208	2	US-08-935-886-8	Sequence 8, Appl
34	46.5	36.0	211	2	US-08-935-886-10	Sequence 10, Appl
35	46	35.7	169	3	US-08-928-941D-35	Sequence 35, Appl
36	46	35.7	169	3	US-09-280-590A-45	Sequence 45, Appl
37	46	35.7	169	4	US-09-892-398-45	Sequence 45, Appl
38	46	35.7	246	4	US-09-248-796A-14865	Sequence 14865, A
39	46	35.7	345	3	US-09-291-922-14	Sequence 14, Appl
40	46	35.7	372	3	US-08-928-941D-16	Sequence 16, Appl
41	46	35.7	372	3	US-09-280-590A-16	Sequence 16, Appl
42	46	35.7	372	4	US-09-892-398-16	Sequence 16, Appl
43	46	35.7	760	3	US-08-928-941D-29	Sequence 29, Appl
44	46	35.7	760	3	US-09-280-590A-29	Sequence 29, Appl
45	46	35.7	760	4	US-09-892-398-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-08-504-538A-12
; Sequence 12, Application US/08504538A
; Patent No. 6004746
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: McCoy, John M.
; APPLICANT: Jessen, Timm H.
; TITLE OF INVENTION: INTERACTION TRAP SYSTEMS FOR DETECTING
; TITLE OF INVENTION: PROTEIN INTERACTIONS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2214
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,538A
; FILING DATE: 07/20/95
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/278,082
; FILING DATE: 07/20/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/259001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 428-0200
; TELEFAX: (617) 428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-504-538A-12

Query Match 100.0%; Score 129; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 QVWSLWALGWRWLRYYGWNM 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 QVWSLWALGWRWLRYYGWNM 20

RESULT 2
US-08-630-052-12
; Sequence 12, Application US/08630052
; Patent No. 6399296
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: McCoy, John M.
; APPLICANT: Jessen, Timm H.
; APPLICANT: Xu, Chanking Wilson
; TITLE OF INVENTION: INTERACTION TRAP SYSTEMS FOR DETECTING PROTEIN
; INTERACTIONS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,052
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/504,538
; FILING DATE: July 20, 1995
; APPLICATION NUMBER: 08/278,082
; FILING DATE: July 20, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Karen F. Lech
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/311001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
US-08-630-052-12

Query Match 100.0%; Score 129; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVWSLWALGWRWLRYYGWNM 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 QVWSLWALGWRWLRYYGWNM 20

RESULT 3
PCT-US95-09307-12
; Sequence 12, Application PC/TUS9509307
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: McCoy, John M.
; APPLICANT: Jessen, Timm H.
; APPLICANT: Xu, Chanking Wilson
; TITLE OF INVENTION: INTERACTION TRAP SYSTEMS FOR
; DETECTING PROTEIN INTERACTIONS

```

```

; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09307
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/288001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US95-09307-12

Query Match 100.0%; Score 129; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVWSLWALGWRWLRYYGWNM 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 QVWSLWALGWRWLRYYGWNM 20

RESULT 4
US-09-252-991A-25051
; Sequence 25051, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25051
; LENGTH: 1086
; TYPE: PET
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25051

Query Match 43.4%; Score 56; DB 4; Length 1086;
Best Local Similarity 62.5%; Pred. No. 43;
Matches 10; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 3 WSLWALGWRWLRYYGW 18

```


APPLICANT: COX, WILLIAM I.

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,587
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/816,155
FILING DATE: 12-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: KOWALSKI, THOMAS J.
REGISTRATION NUMBER: 32,147
REFERENCE/DOCKET NUMBER: 454310-2990
TELEPHONE: 212-588-0800
TELEFAX: 212-588-0500
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 550 amino acids
TYPE: amino acid
STRANDEDNESS: n/a
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-09-079-587-44

Query Match 41.1%; Score 53; DB 3; Length 550;
Best Local Similarity 63.6%; Pred. No. 51;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 WALGWRWLRRY 16
DB 526 WRGWRWLHQY 536

RESULT 9
US-09-560-385A-10
Sequence 10, Application US/09560385A
Patent No. 6703363
GENERAL INFORMATION:
APPLICANT: Boutand, Ariel
TITLE OF INVENTION: Recombinant Laminin 5
FILE REFERENCE: 99-274-C
CURRENT APPLICATION NUMBER: US/09/560,385A
CURRENT FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 1725
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-560-385A-10

Query Match 40.3%; Score 52; DB 4; Length 1725;
Best Local Similarity 54.5%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 WSLWALGWRWL 13
DB 7 WSAWCTGWLWI 17

RESULT 10
US-09-252-991A-31470
Sequence 31470, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31470
LENGTH: 480
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31470

Query Match 39.5%; Score 51; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LGWRWLRR 15
DB 255 LGWRWLRR 262

RESULT 11
US-09-252-991A-31085
Sequence 31085, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31085
LENGTH: 396
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31085

Query Match 39.1%; Score 50.5; DB 4; Length 396;
Best Local Similarity 42.9%; Pred. No. 75;
Matches 9; Conservative 1; Mismatches 4; Indels 7; Gaps 1;

QY 5 LWALG-----WRWLRRYGV 18
DB 115 VWAGVVFVAFQEWGLGRIGW 135

RESULT 12
US-09-328-352-6190
Sequence 6190, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6190
LENGTH: 266
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-6190

Query Match 38.8%; Score 50; DB 4; Length 266;
Best Local Similarity 45.0%; Pred. No. 59;
Matches 9; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

QY 3 WSLWALG-----WRWLRRYGV 18

48 WSMWRTTRRLWWRWRLRW 67

Db

RESULT 13

US-08-637-759B-298
; Sequence 298, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 298:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-637-759B-298

Query Match 38.0%; Score 49; DB 2; Length 61;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 WSLWALGWRWRLRY 16

Db 16 WELLSLNRWYRFR 29

RESULT 14

US-08-871-355A-298
; Sequence 298, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA

ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 298:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-871-355A-298

Query Match 38.0%; Score 49; DB 3; Length 61;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 WSLWALGWRWRLRY 16

Db 16 WELLSLNRWYRFR 29

RESULT 15

US-09-201-945-298
; Sequence 298, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/637,759
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 298:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-201-945-298

Query Match 38.0%; Score 49; DB 3; Length 61;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 WSLWALGWEWLRV 16
Db 16 WELLSLNWYPRF 29

Search completed: October 26, 2004, 06:47:56
Job time : 23 secs

Db 924 QAWKTPNGKWKILYMTLLPLLTQIWMETGWN 959

RESULT 12

CS3374

type IV prepilin peptidase (EC 3.4.99.-) pild - Neisseria lactamica (strain LNP411) (fr
N/Contains: type IV pilin N-methyltransferase (EC 2.1.1.-)

C:Species: Neisseria lactamica
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 29-Jan-1999
C:Accession: CS3374

R;Dupuy, B.; Pugsley, A.P.

J. Bacteriol. 176, 1323-1331, 1994

A:Title: Type IV prepilin peptidase gene of Neisseria gonorrhoeae MS11: presence of a re
A:Reference number: A53374; PMID:94156836; PMID:7906688

A:Accession: CS3374

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A:Molecule type: DNA

A:Residues: 1-102 <DUP>

C:Genetics:

A:Gene: pild

C:Superfamily: type IV prepilin peptidase

C:Keywords: hydrolase; methyltransferase; S-adenosylmethionine

Query Match 38.8%; Score 50; DB 2; Length 102;

Best Local Similarity 54.5%; Pred. No. 10;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 WSLWALGWRWL 13

||| ||| |||

Db 17 WSPWQYGSWI 27

RESULT 13

AD2430

hypothetical protein alr4996 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Name: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: AD2430

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; PMID:21595285; PMID:11759840

A:Accession: AD2430

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-167 <KUR>

A:Cross-references: UNIPROT:Q8YMD8; GB:BA000019; PIDN:BA876695.1; PID:gl71341134; GSPDB:G

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr4996

Query Match 38.8%; Score 50; DB 2; Length 167;

Best Local Similarity 63.6%; Pred. No. 16;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 9 GWRWLRRYGWN 19

||| ||| |||

Db 101 GFYWLSQYGN 111

RESULT 14

AB3647

melittin resistance protein pqab [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C:Accession: AB3647

R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens

A:Reference number: AD3252; PMID:11756688

A:Accession: AB3647

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-477 <KUR>

A:Cross-references: UNIPROT:Q8Y802; GB:AE008918; PIDN:AALS4341.1; PID:gl7985322; GSPDB:G

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEII1099

A:Map position: II

Query Match 38.8%; Score 50; DB 2; Length 477;

Best Local Similarity 47.1%; Pred. No. 44;

Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 2 VWSLWALGWRWLRRYGW 18

||| ||| |||

Db 187 VWMLWRGNWRDLKHLFW 203

RESULT 15

B70584

phosphate transport system permease protein A-1 - Mycobacterium tuberculosis (strain H37Rv)

N:Alternate names: pstA protein

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: B70584; S78423; S42804

R;Cole, S.T.; Broch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: B70584

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-305 <COL>

A:Cross-references: UNIPROT:O86345; GB:Z95209; GB:AL123456; NID:G3261750; PIDN:CAB08490

A:Experimental source: strain H37Rv

R;Braibant, M.; De Wit, L.; Peirs, P.; Kalai, M.; Ooms, J.; Drowart, A.; Huygen, K.; Con

submitted to the EMBL Data library, December 1996

A:Description: Structure of the Mycobacterium tuberculosis Antigen 88, a protein related

A:Reference number: S78423

A:Accession: S78423

A:Molecule type: DNA

A:Residues: 1-4, 'A', '6-207, 'A', '209-304 <BRA>

A:Cross-references: EMBL:X75297; NID:gl182012; PIDN:CAA53045.1; PID:gl182013

A:Note: this is a revision to the sequence from reference S42804

R:Content, J.

submitted to the EMBL Data Library, September 1993

A:Reference number: S42804

A:Accession: S42804

A:Molecule type: DNA

A:Residues: 'MTVLPIGRSGVVAASMLGLRGLGETAVL', 'VILRSAGAGDILVAVRRRLYVRFQDLRCF', 'RIQRTAADR

A:Cross-references: EMBL:X75297

A:Note: this sequence has been revised in reference S78423

C:Genetics:

A:Gene: pstA1

A:Start codon: GTG

C:Superfamily: phoW protein

C:Keywords: phosphate transport

Query Match 38.4%; Score 49.5; DB 2; Length 305;

Best Local Similarity 50.0%; Pred. No. 33;

Matches 10; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

Qy 2 VWSLW---ALGWRWLRRYGW 18

||| ||| |||

Db 49 VLLLVVVIARGWFAVTRSGW 68

Search completed: October 26, 2004, 06:46:22

Job time : 18 secs

us-10-066-965a-1.rpr

Tue Oct 26 08:02:15 2004

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:09:56 ; Search time 93 Seconds
(without alignments)
123.735 Million cell updates/sec

Title: US-10-066-965A-1

Perfect score: 129

Sequence: 1 QVSLWALGWRWLRYYGNM 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	53.5	258	Q55011	streptomyces
2	69	47.3	162	Q8IHB2	drosohila
3	60	46.5	156	Q8S3T1	oryza sativ
4	60	46.5	543	Q943X8	arabidopsis
5	60	46.5	2143	Q9C8A6	arabidopsis
6	60	46.5	2151	Q8RVL2	arabidopsis
7	60	46.5	2159	Q8RUQ1	zea mays (m
8	60	46.5	2159	Q8RVL1	zea mays (m
9	60	46.5	2162	Q6ZF24	oryza sativ
10	60	46.5	2162	BAD07761	oryza sat
11	59.5	46.1	164	Q8BZ40	mus musculus
12	59	45.7	312	Q73VR3	mycobacteri
13	59	45.7	312	AAS05266	mycobacte
14	59	45.7	843	Q8L7W8	arabidopsis
15	59	45.7	847	Q9SVZ1	arabidopsis
16	58	45.0	485	P72844	synchocyst
17	57	44.2	395	Q98414	ovine lentil
18	56	43.4	466	Q7UWQ5	rhodospirell
19	56	43.4	2162	Q8RYA5	oryza sativ
20	54.5	42.2	488	Q6JN63	xanthomonas
21	54.5	42.2	488	AAT39386	xanthomon
22	54.5	42.2	508	Q6PPY9	magnetite-c
23	54.5	42.2	508	AAT01653	magnetite
24	54	41.9	339	Q8PFQ8	xanthomonas
25	54	41.9	584	Q8SU93	encephalito
26	54	41.9	702	Q8C3X8	mus musculus
27	53.5	41.5	228	Q7VDE4	prochloroco
28	53.5	41.5	513	Q7V4C8	prochloroco
29	53	41.1	241	Q933R0	pseudomonas
30	53	41.1	258	Q9HXI6	pseudomonas
31	53	41.1	351	Q7NUL1	chromobacte

RESULT 1

Q55011 PRELIMINARY; PRT; 258 AA.
AC Q55011; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Aminoglycoside-O-phosphotransferase type 5.
GN Name-aph;
OS Streptomyces rimosus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1927;
RN [1] SEQUENCE FROM N.A.
RP Akopiants K.E.; Danilenko V.N.;
RA Akopiants K.E.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2] SEQUENCE FROM N.A.
RA Akopiants K.E.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U24442; AAG03856.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR002575; APH trans.
DR InterPro; IPR011009; Kinase like.
DR Pfam; PF01636; APH; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 258 AA; 28490 MW; D2AD3F681470264E CRC64;

Query Match 53.5%; Score 69; DB 2; Length 258;

Best Local Similarity 56.2%; Pred. No. 0.37; Mismatches 6; Indels 0; Gaps 0;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 WSLWALGWRWLRYYGN 18

Db 56 WQLWGFMAWVRLSGW 71

RESULT 2

Q8IHB2 PRELIMINARY; PRT; 162 AA.
AC Q8IHB2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE AT17081P.
GN Name-CG30266; CG30271; ORFNames-CG30271;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;

```

RN  SEQUENCE FROM N.A.
RP  Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,
RA  Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
RA  George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA  Miranda A., Mungall C.J., Nunoo J., Pacieb J., Paragas V., Park S.,
RA  Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA  Celniker S.,
RL  Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BT001327; AAN71082.1; -.
DR  FlyBase; FBgn050271; CG30271.
SQ  SEQUENCE 162 AA; 20300 MW; 927D2F2060330C09 CRC64;

Query Match 47.3%; Score 61; DB 2; Length 162;
Best Local Similarity 45.5%; Pred. No. 2.6;
Matches 10; Conservative 3; Mismatches 5; Indels 4; Gaps 1;

QY  2 VWSLWALGW----RWLRRYGN 19
DB  51 MWSMWSRRNRNRNRNRNRN 72

RESULT 3
Q8S3T1 PRELIMINARY; PRT; 156 AA.
ID Q8S3T1
AC Q8S3T1
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Putative n-calpain-1 large subunit.
GN Name=49p11.1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Park Y.-J., Rostoks N., Ramakrishna W., SanMiguel P., Shiloff B.,
RA Ma J., Jiang Z., Kleinhofs A., Bennetzen J.,
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF480496; AAL87147.1; -.
DR Gramene; Q8S3T1; -.
SQ SEQUENCE 156 AA; 18241 MW; 0E8A04B23AA15BA6 CRC64;

Query Match 46.5%; Score 60; DB 2; Length 156;
Best Local Similarity 56.2%; Pred. No. 3.4;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY  3 WSLWALGWRWLRYYGW 18
DB  29 WILWAVNRPWRLYSW 44

RESULT 4
Q949X8 PRELIMINARY; PRT; 543 AA.
ID Q949X8
AC Q949X8
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative n-calpain-1 large subunit (Fragment).
GN Name=Atlg5350;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,

```

```

RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.,
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY050822; AAK92757.1; -.
DR NON_TER 543
SQ SEQUENCE 543 AA; 59467 MW; FB988FE03FD3AD48 CRC64;

Query Match 46.5%; Score 60; DB 2; Length 543;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY  3 WSLWALGWRWLRYYGW 18
DB  28 WILWAVNRPWRLYSW 43

RESULT 5
Q9C8A6 PRELIMINARY; PRT; 2143 AA.
ID Q9C8A6
AC Q9C8A6
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE N-calpain-1 large subunit, putative; 13921-23959.
GN Name=F7A10.23;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC027034; AAG51565.1; -.
DR PIR; G96595; G96595.
DR HSSP; P17655; 1KFX.
DR MEROPS; C02.019; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004198; F:calpain activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001300; Peptidase C2.
DR InterPro; IPR000169; Pept_cys_acsite.
DR Pfam; PF01067; Calpain.III; 1.
DR Pfam; PF00648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain.III; 1.
DR SMART; SM00230; Cyspc; 1.
DR PROSITE; PS00203; CALPAIN CAT; 1.
DR PROSITE; PS00139; THIOLEPROTEASE CYS; 1.
SQ SEQUENCE 2143 AA; 236779 MW; B91A15081AF26EC2 CRC64;

Query Match 46.5%; Score 60; DB 2; Length 2143;
Best Local Similarity 56.2%; Pred. No. 41;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY  3 WSLWALGWRWLRYYGW 18
DB  28 WILWAVNRPWRLYSW 43

RESULT 6
Q8RVL2 PRELIMINARY; PRT; 2151 AA.
ID Q8RVL2
AC Q8RVL2;

```

DR	GO:	GO:0005622;	C:intracellular;	IEA.
DR	GO:	GO:0004198;	F:calpain activity;	IEA.
DR	GO:	GO:0006508;	P:proteolysis and peptidolysis;	IEA.
DR	InterPro:	IPR008985;	ConA like lec_gl.	
DR	InterPro:	IPR002052;	N6.Mtase.	
DR	InterPro:	IPR001300;	Peptidase_C2.	
DR	InterPro:	IPR000169;	Pept_cys_acsite.	
DR	Pfam:	PF01067;	Calpain_III; 1.	
DR	Pfam:	PF00648;	Peptidase_C2; 1.	
DR	SMART:	SM00720;	calpain_III; 1.	
DR	SMART:	SM00230;	CysPc; 1.	
DR	PROSITE:	PS50203;	CALPAIN CAT; 1.	
DR	PROSITE:	PS00092;	N6_MTASE; UNKNOWN 1.	
DR	PROSITE:	PS00139;	THIOL_PROTEASE_CYS; 1.	
SQ	SEQUENCE	2159 AA;	238984 MW; 97ICIA709BASFS5D4 CRC64;	
Query Match		46.5%;	Score 60; DB 2; Length 2159;	
Best Local Similarity		56.2%;	Pred. No. 41;	
Matches	9;	Conservative	1; Mismatches 6; Indels 0;	Ga
QY	3	WSLWALGWRRLRYGW	18	
DB	29	KVLWAYNWPRLYSW	44	
RESULT 8				
Q8RVLL	PRELIMINARY; PRT; 2159 AA.			
ID	Q8RVLL			
AC	Q8RVLL			
DT	01-JUN-2002	(TRENBLrel. 21, Created)		
DT	01-JUN-2002	(TRENBLrel. 21, Last sequence update)		
DT	01-MAR-2004	(TRENBLrel. 26, Last annotation update)		
DE	Calpain-like protein.			
GN	Name=dekl			
OS	Zea mays (Maize)			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OX	PACCAD clade; Panicoideae; Andropogoneae; Zea.			
NCBI_TaxID=4577;	[1]			
SEQUENCE FROM N.A.				
RP	TISSUE=developing endosperm;			
RC	PubMed=11929961;			
RX	Lid S.E.; Gruis D., Jung R., Lorentzen J.A., Ananiev E.,			
RA	Chamberlin M., Niu X., Meeley R., Nichols S., Olsen O.A.;			
RT	"The defective kernel 1 (dekl) gene required for aleurone cell			
RT	development in the endosperm of maize grains encodes a membrane			
RT	protein of the calpain gene superfamily."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:5460-5465 (2002).			
DR	EMBL; AY061806; AAL38189.1; -.			
DR	HSP; P17655; I.KFX.			
DR	MEROPS; C02.019; -.			
DR	GO:	GO:0005622;	C:intracellular;	IEA.
DR	GO:	GO:0004198;	F:calpain activity;	IEA.
DR	GO:	GO:0006508;	P:proteolysis and peptidolysis;	IEA.
DR	InterPro:	IPR008985;	ConA like lec_gl.	
DR	InterPro:	IPR002052;	N6.Mtase.	
DR	InterPro:	IPR001300;	Peptidase_C2.	
DR	InterPro:	IPR000169;	Pept_cys_acsite.	
DR	Pfam:	PF01067;	Calpain_III; 1.	
DR	Pfam:	PF00648;	Peptidase_C2; 1.	
DR	SMART:	SM00720;	calpain_III; 1.	
DR	SMART:	SM00230;	CysPc; 1.	
DR	PROSITE:	PS50203;	CALPAIN CAT; 1.	
DR	PROSITE:	PS00092;	N6_MTASE; UNKNOWN 1.	
DR	PROSITE:	PS00139;	THIOL_PROTEASE_CYS; 1.	
SQ	SEQUENCE	2159 AA;	238997 MW; 9F995D3F9EAC1315 CRC64;	
Query Match		46.5%;	Score 60; DB 2; Length 2159;	
Best Local Similarity		56.2%;	Pred. No. 41;	
Matches	9;	Conservative	1; Mismatches 6; Indels 0;	Ga
QV	3	WSLWALGWRRLRYGW	18	

Db 29 WVLWVNNRWPRLYSW 44

RESULT 9

Q6ZF24 ID Q6ZF24 PRELIMINARY; PRT; 2162 AA.

AC Q6ZF24

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Dek1-calpain-like protein.

GN Name-OJ1311_H06.4;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoidae; Oryzeae; Oryza.

NCBI_TaxID=39947;

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RA Sasaki T., Matsumoto T., Yamamoto K.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP004161; BAD07761.1;

DR InterPro; IPR008985; ConA_like_1ec_g1.

DR InterPro; IPR002052; N6_Mtase.

DR InterPro; IPR001300; Peptidase_C2.

DR InterPro; IPR000169; Pept_cys_acsite.

DR Pfam; PF01067; Calpain_III; 1.

DR Pfam; PF00648; Peptidase_C2; 1.

DR PRINTS; PR00704; CALPAIN.

DR SMART; SM00720; calpain_III; 1.

DR SMART; SM00230; CysPc; 1.

DR PROSITE; PS00203; CALPAIN_CAT; 1.

DR PROSITE; PS00092; N6_MTASE; UNKNOWN 1.

DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.

DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.

SQ SEQUENCE 2162 AA; 239856 MW; 9FF448AE2B23F0B CRC64;

Query Match 46.5%; Score 60; DB 2; Length 2162;

Best Local Similarity 56.2%; Pred. No. 41;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 WSLWALGWRWRLRYGW 18

Db 29 WILWVNNRWPRLYSW 44

RESULT 10

BAD07761 ID BAD07761 PRELIMINARY; PRT; 2162 AA.

AC BAD07761

DT 02-MAR-2004 (TrEMBLrel. 27, Created)

DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)

DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)

DE Dek1-calpain-like protein.

GN OJ1311_H06.4.

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoidae; Oryzeae; Oryza; Oryza sativa.

NCBI_TaxID=39947;

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=cv. Nipponbare;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT clon:OJ1311_H06.4;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP004161; BAD07761.1;

SQ SEQUENCE 2162 AA; 239856 MW; 9FF448AE2B23F0B CRC64;

Query Match 46.5%; Score 60; DB 2; Length 2162;

Best Local Similarity 56.2%; Pred. No. 41;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 WSLWALGWRWRLRYGW 18

Db 29 WILWVNNRWPRLYSW 44

RESULT 11

Q8BZ40 ID Q8BZ40 PRELIMINARY; PRT; 164 AA.

AC Q8BZ40

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Mus musculus adult male bone cDNA, RIKEN full-length enriched library;

DE clone:9830169C09 product:ALANYL-TRNA SYNTHETASE (EC 6.1.1.7) (ALANINE-TRNA LIGASE) (ALARS) homolog.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=C57BL/6J; TISSUE=Bone;

RX MEDLINE=99279253; PubMed=10349636;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44(1999).

RN [2]

RP SEQUENCE FROM N.A.

RA STRAIN=C57BL/6J; TISSUE=Bone;

RX MEDLINE=21085660; PubMed=11217851;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [3]

RP SEQUENCE FROM N.A.

RA STRAIN=C57BL/6J; TISSUE=Bone;

RT The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [4]

RP SEQUENCE FROM N.A.

RA STRAIN=C57BL/6J; TISSUE=Bone;

RX MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,

RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

RN [5]

RP SEQUENCE FROM N.A.

RA STRAIN=C57BL/6J; TISSUE=Bone;

RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,

RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,

RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,

RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).

RN [6]

RP SEQUENCE FROM N.A.

RA STRAIN=C57BL/6J; TISSUE=Bone;

RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,

RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

RA Hayashida K., Hayatsu M., Hiramoto K., Hiraoka T., Hirozawa T.,

RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,

RA Katoh H., Kawai J., Kotani Y., Kondo S., Kono H., Kouda M., Koya S.,

```

RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Nunazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK036738; BAC29558.1; -.
DR GO; GO:0004813; P:alanine-tRNA ligase activity; IEA.
DR GO; GO:0016874; P:ligase activity; IEA.
KW Aminoacyl-tRNA synthetase; Ligase.
SQ SEQUENCE 164 AA; 18827 MW; D73DC68F0B60F610 CRC64;

Query Match 46.1%; Score 59.5; DB 2; Length 164;
Best Local Similarity 58.8%; Pred. No. 4.2; Mismatches 2; Indels 1; Gaps 1;
Matches 10; Conservative 2;

QY 3 WSLWAL-GWRWLRYYGW 18
Db 75 WTFWALAGFRWLPFGMW 91
||| ||| ||| ||| |||
|: ||| |: ||| |||

RESULT 12
QY3VR3 PRELIMINARY; PRT; 312 AA.
AC Q73VR3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP2949c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017237; AAS05266.1; -.
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; Plsc; 1.
KW SMART proteome; Hypothetical protein.
SQ SEQUENCE 312 AA; 35001 MW; 577E4AE35A177103 CRC64;

Query Match 45.7%; Score 59; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 8.9; Mismatches 3; Indels 0; Gaps 0;
Matches 8; Conservative 3;

QY 2 VWSLWALGWRWLRYYG 17
Db 101 VWDATVGLQWRRFG 116
||| ||| ||| ||| |||
|: ||| |: ||| ||| |||

RESULT 13
AAS05266 PRELIMINARY; PRT; 312 AA.
AC AAS05266;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN MAP2949C.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K10;

```

```

RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017237; AAS05266.1; -.
KW Hypothetical protein.
SQ SEQUENCE 312 AA; 35001 MW; 577E4AE35A177103 CRC64;

Query Match 45.7%; Score 59; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 8.9; Mismatches 3; Indels 0; Gaps 0;
Matches 8; Conservative 3;

QY 2 VWSLWALGWRWLRYYG 17
Db 101 VWDATVGLQWRRFG 116
||| ||| ||| ||| |||
|: ||| |: ||| ||| |||

RESULT 14
QY8L7W8 PRELIMINARY; PRT; 843 AA.
AC Q8L7W8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AT4934260/F10M10_30.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY125494; AAM78086.1; -.
DR EMBL; BT002722; AAC11638.1; -.
DR InterPro; IPR008928; Glyco trans 6hp.
SQ SEQUENCE 843 AA; 93724 MW; 2EF1315DF98A1365 CRC64;

Query Match 45.7%; Score 59; DB 2; Length 843;
Best Local Similarity 47.1%; Pred. No. 23; Mismatches 6; Indels 0; Gaps 0;
Matches 8; Conservative 3;

QY 2 VWSLWALGWRWLRYYGW 18
Db 499 VVALPFGGAWLCTHAW 515
||| ||| ||| ||| |||
|: ||| |: ||| ||| |||

RESULT 15
QY9YZ1 PRELIMINARY; PRT; 847 AA.
AC Q9SYZ1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein AT4934260.
GN Name=AT4934260;
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OX eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;

RP [1]
 RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AL035521; CAB36703.1; -
 DR EMBL; AL161585; CAB80143.1; -

DR PIR; T04772; T04772
 DR InterPro; IPR008928; Glyco_trans_6hp.

KW Hypothetical protein.

SQ SEQUENCE 847 AA; 94023 MW; 98D95FB21BF5EA5D CRC64;

Query Match 45.7%; Score 59; DB 2; Length 847;
 Best Local Similarity 47.1%; Pred. No. 23;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 VWSLWALGWRWLRYYGW 18
 ||:||:|:|:|
 Db 503 VVALWPMGGAWLCTHAW 519

Search completed: October 26, 2004, 06:39:20
 Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:45:25 ; Search time 64.25 seconds
(without alignments)
100.781 Million cell updates/sec

Title: US-10-066-965A-1

Perfect score: 129

Sequence: 1 QVMSLWALGWRWLRYYGWNW 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	20	14	US-10-162-538-12
2	129	100.0	20	14	US-10-066-965A-1
3	129	100.0	20	14	US-10-066-965A-7
4	117	90.7	20	14	US-10-066-965A-2
5	117	90.7	20	14	US-10-066-965A-8
6	60	46.5	2094	16	US-10-437-963-187850
7	60	46.5	2115	14	US-10-208-948-4
8	60	46.5	2150	14	US-10-208-948-27
9	60	46.5	2157	17	US-10-696-616-16
10	60	46.5	2159	14	US-10-208-948-24
11	59	45.7	73	9	US-09-764-860-479
12	59	45.7	73	14	US-10-074-095-479
13	59	45.7	73	14	US-10-212-872-479

14 45.7 135 15 US-10-424-599-149236 Sequence 149236,
15 43.4 2160 16 US-10-437-963-187848 Sequence 187848,
16 43.0 15 US-10-425-134-44142 Sequence 44142, A
17 42.6 485 15 US-10-424-599-189763 Sequence 189763, A
18 42.6 486 15 US-10-425-114-44744 Sequence 44744, A
19 42.6 511 15 US-10-424-599-203054 Sequence 203054, A
20 41.1 96 16 US-10-437-963-142534 Sequence 142534,
21 41.1 258 9 US-09-815-242-11957 Sequence 11957, A
22 41.1 258 15 US-10-282-122A-66556 Sequence 66556, A
23 41.1 411 14 US-10-369-493-5089 Sequence 5089, Ap
24 41.1 712 16 US-10-437-963-144606 Sequence 144606,
25 40.7 830 16 US-10-437-963-198121 Sequence 198121,
26 40.7 840 16 US-10-437-963-198124 Sequence 198124,
27 40.3 105 14 US-10-029-386-32277 Sequence 32277, A
28 40.3 358 15 US-10-264-237-2747 Sequence 2747, Ap
29 40.3 1725 15 US-10-603-725-10 Sequence 10, Appl
30 39.9 408 16 US-10-437-963-153557 Sequence 153557,
31 39.5 68 9 US-09-796-692-721 Sequence 721, App
32 39.5 68 14 US-10-040-862-721 Sequence 721, App
33 39.5 68 15 US-10-057-475B-721 Sequence 721, App
34 39.5 68 15 US-10-154-884B-721 Sequence 721, App
35 39.5 68 16 US-10-764-324-721 Sequence 721, App
36 39.5 261 15 US-10-282-122A-50652 Sequence 50652, A
37 39.1 73 16 US-10-767-701-54428 Sequence 54428, A
38 39.1 104 16 US-10-437-963-106350 Sequence 106350,
39 39.1 300 15 US-10-282-122A-49338 Sequence 49338, A
40 39.1 538 15 US-10-282-122A-49335 Sequence 49335, A
41 38.8 219 16 US-10-767-701-52070 Sequence 52070, A
42 38.4 139 15 US-10-425-114-52367 Sequence 15413, A
43 38.4 798 14 US-10-369-493-15413 Sequence 15413, A
44 38.4 798 14 US-10-369-493-15781 Sequence 15781, A
45 38.4 798 14 US-10-369-493-16163 Sequence 16163, A

ALIGNMENTS

RESULT 1

US-10-162-538-12
; Sequence 12, Application US/10162538
; Publication No. US20030113749A1
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; McCoy, John M.
; Jessen, Timm H.
; Xu, Channing Wilson
; TITLE OF INVENTION: INTERACTION TRAP SYSTEMS FOR DETECTING
; PROTEIN
; INTERACTIONS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/10/162,538
; FILING DATE: 04-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER: US/08/630,052
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/504,538
; FILING DATE: July 20, 1995
; APPLICATION NUMBER: 08/278,082
; FILING DATE: July 20, 1994

ATTORNEY/AGENT INFORMATION:
NAME: Karen F. Lech
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/311001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 20
TYPE: amino acid
STRANDEDNESS: No. US20030113749A1 Relevant
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-162-538-12

Query Match 100.0%; Score 129; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVSWLWALGWRWLRYYGWNM 20
DB 1 QVSWLWALGWRWLRYYGWNM 20

RESULT 2
US-10-066-965A-1
; Sequence 1, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-1

Query Match 100.0%; Score 129; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVSWLWALGWRWLRYYGWNM 20
DB 1 QVSWLWALGWRWLRYYGWNM 20

RESULT 3
US-10-066-965A-7
; Sequence 7, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7

LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-7

Query Match 100.0%; Score 129; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVSWLWALGWRWLRYYGWNM 20
DB 1 QVSWLWALGWRWLRYYGWNM 20

RESULT 4
US-10-066-965A-2
; Sequence 2, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-2

Query Match 90.7%; Score 117; DB 14; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.2e-06;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QVSWLWALGWRWLRYYGWNM 20
DB 1 QVSWLWALGWRWLRYYGWNM 20

RESULT 5
US-10-066-965A-8
; Sequence 8, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-8

Query Match 90.7%; Score 117; DB 14; Length 20;

```
Best Local Similarity 90.0%; Pred. No. 1.2e-06;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QVSWLWALGWRWLRYYGNM 20
Db 1 QVSSWALGWRWLRYYGWM 20

RESULT 6
US-10-437-963-187850
; Sequence 187850, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(33221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 187850
; LENGTH: 2094
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84512C.1.pep
US-10-437-963-187850

Query Match 46.5%; Score 60; DB 16; Length 2094;
Best Local Similarity 56.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 WSLWALGWRWLRYYGW 18
Db 29 WILWAVNWRPWRLYSW 44

RESULT 7
US-10-208-948-4
; Sequence 4, Application US/10208948
; Publication No. US20030074699A1
; GENERAL INFORMATION:
; APPLICANT: Tarczynski, Mitchell C.
; APPLICANT: Olsen, Odd-Arne
; APPLICANT: Shen, Bo
; APPLICANT: Lid, Stein E.
; APPLICANT: Li, Changjiang
; APPLICANT: Jung, Rudolf
; APPLICANT: Gruis, Darren B.
; APPLICANT: Lorentzen, Jennifer A.
; APPLICANT: Ananiev, Evgueni
; APPLICANT: Nichols, Scott E.
; APPLICANT: Wang, Cunxi
; TITLE OF INVENTION: Methods for Improving Seed and Grain
; TITLE OF INVENTION: Characteristics
; FILE REFERENCE: 1390
; CURRENT APPLICATION NUMBER: US/10/208,948
; CURRENT FILING DATE: 2002-07-30
; PRIOR APPLICATION NUMBER: 60/309,719
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/337,444
; PRIOR FILING DATE: 2001-10-25
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2115
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-208-948-4

Query Match 46.5%; Score 60; DB 14; Length 2150;
Best Local Similarity 56.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 WSLWALGWRWLRYYGW 18
Db 28 WILWAVNWRPWRLYSW 43

RESULT 9
US-10-696-616-16
; Sequence 16, Application US/10696616
; Publication No. US20040191850A1
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua
; APPLICANT: Tossberg, John
; APPLICANT: Zhou, Qing
; APPLICANT: McElver, John
; APPLICANT: Aux, George
; APPLICANT: Budziszewski, Greg
; APPLICANT: Thomas, Carla
; APPLICANT: Patton, David
; APPLICANT: Frye, Catherine
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Essential for Plant Growth
; TITLE OF INVENTION: Development and Uses Thereof
; FILE REFERENCE: 70166USNP
; CURRENT APPLICATION NUMBER: US/10/696,616
```

```
; TYPE: PRT
; ORGANISM: arabidopsis
US-10-208-948-4

Query Match 46.5%; Score 60; DB 14; Length 2115;
Best Local Similarity 56.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 WSLWALGWRWLRYYGW 18
Db 28 WILWAVNWRPWRLYSW 43

RESULT 8
US-10-208-948-27
; Sequence 27, Application US/10208948
; Publication No. US20030074699A1
; GENERAL INFORMATION:
; APPLICANT: Tarczynski, Mitchell C.
; APPLICANT: Olsen, Odd-Arne
; APPLICANT: Shen, Bo
; APPLICANT: Lid, Stein E.
; APPLICANT: Li, Changjiang
; APPLICANT: Jung, Rudolf
; APPLICANT: Gruis, Darren B.
; APPLICANT: Lorentzen, Jennifer A.
; APPLICANT: Ananiev, Evgueni
; APPLICANT: Nichols, Scott E.
; APPLICANT: Wang, Cunxi
; TITLE OF INVENTION: Methods for Improving Seed and Grain
; TITLE OF INVENTION: Characteristics
; FILE REFERENCE: 1390
; CURRENT APPLICATION NUMBER: US/10/208,948
; CURRENT FILING DATE: 2002-07-30
; PRIOR APPLICATION NUMBER: 60/309,719
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/337,444
; PRIOR FILING DATE: 2001-10-25
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 2150
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-208-948-27

Query Match 46.5%; Score 60; DB 14; Length 2150;
Best Local Similarity 56.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 WSLWALGWRWLRYYGW 18
Db 28 WILWAVNWRPWRLYSW 43

RESULT 9
US-10-696-616-16
; Sequence 16, Application US/10696616
; Publication No. US20040191850A1
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua
; APPLICANT: Tossberg, John
; APPLICANT: Zhou, Qing
; APPLICANT: McElver, John
; APPLICANT: Aux, George
; APPLICANT: Budziszewski, Greg
; APPLICANT: Thomas, Carla
; APPLICANT: Patton, David
; APPLICANT: Frye, Catherine
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Essential for Plant Growth
; TITLE OF INVENTION: Development and Uses Thereof
; FILE REFERENCE: 70166USNP
; CURRENT APPLICATION NUMBER: US/10/696,616
```

```

/ CURRENT FILING DATE: 2003-10-29
/ PRIOR APPLICATION NUMBER: US 60/423,519
/ PRIOR FILING DATE: 2002-11-04
/ NUMBER OF SEQ ID NOS: 73
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 16
/ LENGTH: 2157
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (105)..(105)
/ OTHER INFORMATION: The 'Xaa' at location 105 stands for Ala, or Thr.
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (516)..(516)
/ OTHER INFORMATION: The 'Xaa' at location 516 stands for Glu, Ala, Gln, or Pro.
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (553)..(553)
/ OTHER INFORMATION: The 'Xaa' at location 553 stands for Leu.
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1659)..(1659)
/ OTHER INFORMATION: Xaa= any amino acid
/ US-10-696-616-16

```

; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244

```
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
```

```
Query Match 45.7%; Score 59; DB 14; Length 73;
Best Local Similarity 47.1%; Pred. No. 11;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 2 VWSLWALGWRWLRRYGW 18
Db 16 IWSMWLPGEQWLXPXSW 32
```

RESULT 13

```
US-10-212-872-479
; Sequence 479, Application US/10212872
; Publication No. US20030215893A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C2
; CURRENT APPLICATION NUMBER: US/10/212,872
; CURRENT FILING DATE: 2002-08-07
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 479
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (70)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

```
US-10-212-872-479
```

```
Query Match 45.7%; Score 59; DB 14; Length 73;
Best Local Similarity 47.1%; Pred. No. 11;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 2 VWSLWALGWRWLRRYGW 18
Db 16 IWSMWLPGEQWLXPXSW 32
```

RESULT 14

```
US-10-424-599-149236
; Sequence 149236, Application US/10424599
; Publication No. US20040031072A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149236
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105784C.1.pep
; US-10-424-599-149236
```

```
Query Match 45.7%; Score 59; DB 15; Length 135;
Best Local Similarity 53.3%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 3 WSLWALGWRWLRRYG 17
Db 118 WRKMGCGWRMGARWG 132
```

RESULT 15

```
US-10-437-963-187848
; Sequence 187848, Application US/10437963
; Publication No. US20040123343A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 187848
; LENGTH: 2160
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(2160)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84510C.1.pep
; US-10-437-963-187848
```

```
Query Match 43.4%; Score 56; DB 16; Length 2160;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 3 WSLWALGWRWLRRYG 18
Db 29 WILWAVNWRPWLPSW 44
```

```
Search completed: October 26, 2004, 07:10:18
Job time : 64.25 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:19:49 ; Search time 17 Seconds
(without alignments)
113.196 Million cell updates/sec

Title: US-10-066-965A-2

Perfect score: 129

Sequence: 1 QVWSSWALGWRWLRYYGWM 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	57	44.2	377	B83454	probable alkane hy
2	55	42.6	123	I48192	surface antigen -
3	55	42.6	282	SAVL64	middle surface ant
4	55	42.6	426	SAVL6	large surface anti
5	55	42.6	428	SAVL5	large surface anti
6	55	42.6	431	SAVL59	large surface anti
7	55	42.6	431	SAVL7	large surface anti
8	55	42.6	431	SAVL8	large surface anti
9	54	41.9	2143	G95595	hypothetical prote
10	53	41.1	67	F72226	hypothetical prote
11	53	41.1	847	T04772	hypothetical prote
12	52	40.3	102	C53374	type IV prepilin p
13	52	40.3	398	F72335	hypothetical prote
14	52	40.3	485	S74708	hypothetical prote
15	51	39.5	135	W4B277	gene 4.7 protein -
16	51	39.5	194	D75078	hypothetical prote
17	51	39.5	297	C83311	hypothetical prote
18	51	39.5	431	SAVL62	large surface anti
19	51	39.5	795	B83755	hypothetical prote
20	50	38.8	226	JQ1579	major surface anti
21	50	38.8	226	JQ1580	major surface anti
22	50	38.8	226	JQ2119	surface antigen -
23	50	38.8	226	JQ2225	small surface prot
24	50	38.8	226	JQ2122	surface antigen -
25	50	38.8	226	JQ2121	surface antigen -
26	50	38.8	226	JQ2120	surface antigen -
27	50	38.8	251	JQ2226	middle surface pro
28	49.5	38.8	400	JQ2230	large surface prot
29		38.4	382	A83325	conserved hypothet

30	48.5	37.6	389	2	AH2931	conserved hypothet
31	48.5	37.6	425	2	F98350	hypothetical prote
32	48	37.2	165	2	H75284	hypothetical prote
33	48	37.2	327	2	H82443	hypothetical prote
34	48	37.2	350	2	T40466	probable acetyltra
35	47.5	36.8	196	2	AE2918	acetyltransferase
36	47.5	36.8	207	2	H97692	conserved hypothet
37	47.5	36.8	225	2	T49769	hypothetical prote
38	47.5	36.8	336	2	D71474	probable muramoyl-
39	47	36.4	71	2	G72226	hypothetical prote
40	47	36.4	258	2	A83169	o-acetylserine syn
41	47	36.4	259	2	AF0999	conserved hypothet
42	47	36.4	268	2	F65134	hypothetical 30.0K
43	47	36.4	274	2	E83215	probable permease
44	47	36.4	278	2	AI1590	phage protein homo
45	47	36.4	330	2	D90862	peptide transport

ALIGNMENTS

RESULT 1

B83454

probable alkane hydroxylase PA1525 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 16-Aug-2004

C;Accession: B83454

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Brj
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathog
A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: B83454

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-377 <STO>

A;Cross-references: UNIPROT:Q9I318; GB:AE004581; GB:AE004091; NID:g9947482; PIDN:AAGC4914
A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA1525

C;Superfamily: Alkane 1-monooxygenase

Query Match 44.2%; Score 57; DB 2; Length 377;
Best Local Similarity 42.9%; Pred. No. 5.2;
Matches 9; Conservative 2; Mismatches 6; Indels 4; Gaps 1;

QY 2 VWSWALG---WRWLRYYGW 18

Db 101 VWSGWLAHETRWVQQLGW 121

RESULT 2

I48192

surface antigen - woodchuck (fragment)

C;Species: Marmota monax (woodchuck)

C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004

C;Accession: I48192

R;Yamazoe, M.; Nakai, S.; Ogasawara, N.; Yoshikawa, H.

Gene 100, 139-146, 1991

A;Title: Integration of woodchuck hepatitis virus (whv) DNA at two chromosomal sites (v-
A;Reference number: I48190; MUID:91276235; PMID:2055466

A;Accession: I48192

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-123 <RES>

A;Cross-references: UNIPROT:Q90137; GB:M60766; NID:gi191471; PIDN:AAA37108.1; PID:gi191472
C;Superfamily: hepatitis B virus surface antigen

C;Keywords: surface antigen

Query Match 42.6%; Score 55; DB 2; Length 123;
Best Local Similarity 57.9%; Pred. No. 3.1;
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 2;

A:Molecule type: DNA

A:Residues: 1-431 <COH>

A:Cross-references: UNIPROT:P12909; GB:M18752; NID:g336136; PIDN:AAA46766.1; PID:g336137

C:Genetics:

A:Gene: pre-S1/pre-S2/S

C:Superfamily: hepatitis B virus surface antigen

C:Keywords: glycoprotein; surface antigen

F:150-431/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>

F:210-431/Product: major surface antigen (gene S) #status predicted <MSA>

F:32,94,152/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 42.6%; Score 55; DB 1; Length 431;
Best Local Similarity 57.9%; Pred. No. 10;
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 2;

Qy 4 SSWALG---WFW-LRRYGV 18

|||||

Db 359 SSWALGNLYWEWALRFSW 377

RESULT 8

SAVLW8

Large surface antigen - woodchuck hepatitis virus (clone 8)

N:Contains: major surface antigen; middle surface antigen

C:Species: woodchuck hepatitis virus

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C:Accession: B32397

R:Girones, R.; Core, P.J.; Hornbuckle, W.E.; Tennant, B.C.; Gerin, J.L.; Purcell, R.H.;

Proc. Natl. Acad. Sci. U.S.A. 86, 1846-1849, 1989

A:Title: Complete nucleotide sequence of a molecular clone of woodchuck hepatitis virus

A:Reference number: A94222; MUID:89184524; PMID:2928306

A:Accession: B32397

A:Molecule type: DNA

A:Residues: 1-431 <GIR>

A:Cross-references: UNIPROT:P17400; GB:J04514

C:Genetics:

A:Gene: pre-S1/pre-S2/S

C:Superfamily: hepatitis B virus surface antigen

C:Keywords: glycoprotein; surface antigen

F:150-431/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>

F:210-431/Product: major surface antigen (gene S) #status predicted <MSA>

F:32,94,152/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 42.6%; Score 55; DB 1; Length 431;
Best Local Similarity 57.9%; Pred. No. 10;
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 2;

Qy 4 SSWALG---WFW-LRRYGV 18

|||||

Db 359 SSWALGNLYWEWALRFSW 377

RESULT 9

G96595

hypothetical protein F7A10.23 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: G96595

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G96595

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2143 <STO>

A:Cross-references: UNIPROT:Q9C8A6; GB:AE005173; NID:g10645420; PIDN:AAG21537.1; GSPDB:G1

C:Genetics:

A:Gene: F7A10.23

A:Map position: 1

Query Match 41.9%; Score 54; DB 2; Length 2143;

Best Local Similarity 50.0%; Pred. No. 64;

Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 3 WSSWALGWRWLRRYGV 18

|||||

Db 28 WILWAVNWRPRLYSW 43

RESULT 10

F72226

hypothetical protein - Thermotoga maritima (strain MSB9)

C:Species: Thermotoga maritima

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C:Accession: F72226

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: F72226

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-67 <ARN>

A:Cross-references: UNIPROT:Q9X208; GB:AE001808; NID:g4982233; PIDN:AAD3674;

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM1680

Query Match 41.1%; Score 53; DB 2; Length 67;

Best Local Similarity 53.3%; Pred. No. 3.1;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 5 SWALGWRWLRRYGV 19

|||||

Db 28 SWARPGWGRGFGY 42

RESULT 11

T04772

hypothetical protein F10M10.30 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T04772

R:Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Hoheisel, J.; Mewes, H.W.; Mayer,

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15384

A:Accession: T04772

A:Molecule type: DNA

A:Residues: 1-847 <BEV>

A:Cross-references: UNIPROT:Q9SVZ1; EMBL:AL035521

A:Experimental source: cultivar Columbia; BAC clone F10M10

C:Genetics:

A:Map position: 4

A:Insertions: 91/1; 139/3; 416/1; 475/3; 528/3; 603/3; 683/1; 748/2

A>Note: F10M10.30

C:Superfamily: Arabidopsis thaliana hypothetical protein F10M10.30

Query Match 41.1%; Score 53; DB 2; Length 847;

Best Local Similarity 41.2%; Pred. No. 35;

Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 VWSSWALGWRWLRRYGV 18

|||||

Db 503 VWALWPGGAWLCTHAW 519

RESULT 12

C53374

type IV prepilin peptidase (EC 3.4.99.-) piliD - Neisseria lactamica (strain LNP411) (fr

N;Contains: type IV pilin N-methyltransferase (EC 2.1.1.-)

C;Species: Neisseria lactamica

C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 29-Jan-1999

C;Accession: C53374

R;Dupuy, B.; Pugsley, A.P.

J. Bacteriol. 176, 1323-1331, 1994

A;Title: Type IV prepilin peptidase gene of Neisseria gonorrhoeae MS11: presence of a re

A;Reference number: A53374; MUID:94156836; PMID:17906688

A;Accession: C53374

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A;Molecule type: DNA

A;Residues: 1-102 <DUP>

C;Genetics:

A;Gene: piliD

C;Superfamily: type IV prepilin peptidase

C;Keywords: hydrolase; methyltransferase; S-adenosylmethionine

Query Match 40.3%; Score 52; DB 2; Length 102;

Best Local Similarity 54.5%; Pred. No. 6.2;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 WSSWALGWRWL 13

||| |||

Db 17 WSPQYGVSWI 27

RESULT 13

F72335

hypothetical protein TM0760 - Thermotoga maritima (strain MSB8)

C;Species: Thermotoga maritima

C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C;Accession: F72335

R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.

C.M.

Nature 399, 323-329, 1999

A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: F72335

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-398 <ARN>

A;Cross-references: UNIPROT:Q9WZL9; GB:AE001746; GB:AE000512; NID:g4981285; PIDN:AAD3584

A;Experimental source: strain MSB8

C;Genetics:

A;Gene: TM0760

Query Match 40.3%; Score 52; DB 2; Length 398;

Best Local Similarity 61.5%; Pred. No. 23;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SWALGWRWLRYYG 17

||| |||

Db 113 WSLGWYVVRKG 125

RESULT 14

S74708

hypothetical protein slr1306 - Synechocystis sp. (strain PCC 6803)

C;Species: Synechocystis sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C;Accession: S74708

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

s.

A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S74708

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-485 <KAN>

A;Cross-references: UNIPROT:P72844; EMBL:D90901; GB:AB001339; NID:gl651897; PIDN:BAAL6859

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 40.3%; Score 52; DB 2; Length 485;

Best Local Similarity 45.0%; Pred. No. 27;

Matches 9; Conservative 1; Mismatches 4; Indels 6; Gaps 1;

Qy 3 WSSWALG-----WRWLRYY 16

||| |||

Db 45 WGLWALSGLGLVIRWRRRH 64

RESULT 15

W4BP77

gene 4.7 protein - phage T7

C;Species: phage T7

C;Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 09-Jul-2004

C;Accession: A04411; S42310

R;Dunn, J.J.; Thompson, K.

submitted to the Nucleic Acid Sequence Database, September 1982

A;Reference number: A94615

A;Accession: A04411

A;Molecule type: DNA

A;Residues: 1-135 <DUN>

A;Cross-references: UNIPROT:P03786

R;Dunn, J.J.; Studier, F.W.

J. Mol. Biol. 166, 477-535, 1983

A;Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 ge

A;Reference number: S42283; MUID:83241725; PMID:6864790

A;Accession: S42310

A;Molecule type: DNA

A;Residues: 1-135 <DUN>

A;Cross-references: EMBL:V01146; NID:g431187; PIDN:CAA24411.1; PID:gl5590

C;Genetics:

A;Gene: 4.7

A;Map position: 34.87-35.88

C;Superfamily: phage T7 gene 4.7 protein

Query Match

39.5%; Score 51; DB 1; Length 135;

Best Local Similarity 63.6%; Pred. No. 11;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 8 LGWRWLRYYG 18

||| |||

Db 38 LGWTWTRQTGW 48

Search completed: October 26, 2004, 06:46:22

Job time : 17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:09:56 ; Search time 93 Seconds
(without alignments)
123.736 Million cell updates/sec

Title: US-10-066-965A-2
Perfect score: 129
Sequence: 1 QVSSWALGWRWLRYYGWM 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	48.8	258	Q55011	Q55011 streptomyces
2	61	47.3	312	Q73VR3	Q73vr3 mycobacteri
3	61	47.3	312	AAS05266	AAS05266 mycobacte
4	60.5	46.9	488	Q6JN63	Q6JN63 xanthomonas
5	60.5	46.9	488	AAT39386	AAT39386 xanthomon
6	58	45.0	466	Q7UWQ5	Q7UWQ5 rhodospirill
7	57.5	44.6	164	Q8BZ40	Q8BZ40 mus musculi
8	57	44.2	163	Q7U9N8	Q7U9N8 synechococc
9	57	44.2	241	Q933R0	Q933R0 pseudomonas
10	57	44.2	377	Q6H927	Q6H927 pseudomonas
11	57	44.2	377	Q6H936	Q6H936 pseudomonas
12	57	44.2	377	Q6H941	Q6H941 pseudomonas
13	57	44.2	377	Q913I8	Q913I8 pseudomonas
14	57	44.2	700	Q6SQM6	Q6SQM6 pseudomonas
15	57	44.2	700	AAR13803	AAR13803 pseudomon
16	55.5	43.0	399	Q8PF08	Q8PF08 xanthomonas
17	55	42.6	123	Q90137	Q90137 marmota mon
18	55	42.6	222	Q83760	Q83760 woodchuck h
19	55	42.6	222	Q89245	Q89245 woodchuck h
20	55	42.6	282	1 VMSA_WHV6	1 VMSA_WHV6
21	55	42.6	426	1 VMSA_WHV1	1 VMSA_WHV1
22	55	42.6	426	2 Q774I8	2 Q774I8 woodchuck h
23	55	42.6	428	1 VMSA_HPBGS	1 VMSA_HPBGS
24	55	42.6	431	1 VMSA_WHV59	1 VMSA_WHV59
25	55	42.6	431	1 VMSA_WHV7	1 VMSA_WHV7
26	55	42.6	431	1 VMSA_WHV8I	1 VMSA_WHV8I
27	54.5	42.2	169	2 Q8H7U5	2 Q8H7U5 oryza sativ
28	54.5	42.2	452	1 CB11_ARATH	1 CB11_ARATH
29	54	41.9	156	2 Q8S3T1	2 Q8S3T1 arabidopsis
30	54	41.9	429	2 Q8AAZ7	2 Q8AAZ7 arabidopsis
31	54	41.9	543	2 Q949X8	2 Q949X8 arabidopsis

32	54	41.9	588	2	Q94E52	Q94e52 oryza sativ
33	54	41.9	1725	2	P70570	P70570 rattus norv
34	54	41.9	2143	2	Q9C8A6	Q9C8A6 arabidopsis
35	54	41.9	2151	2	Q8RVL2	Q8RVL2 arabidopsis
36	54	41.9	2159	2	Q8RUQ1	Q8RUQ1 zea mays (m
37	54	41.9	2159	2	Q8RVL1	Q8RVL1 zea mays (m
38	54	41.9	2162	2	Q6ZFZ4	Q6ZFZ4 oryza sativ
39	54	41.9	2162	2	BAD07761	BAD07761 oryza sat
40	53	41.1	67	2	Q9X208	Q9X208 thermotoga
41	53	41.1	222	2	Q64901	Q64901 arctic grou
42	53	41.1	234	2	Q73TC3	Q73tc3 mycobacteri
43	53	41.1	234	2	AAS06345	AAS06345 mycobacte
44	53	41.1	261	2	Q6IH30	Q6IH30 drosophila
45	53	41.1	282	2	Q64900	Q64900 arctic grou

ALIGNMENTS

RESULT 1
Q55011 PRELIMINARY; PRT; 258 AA.
ID Q55011
AC Q55011;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-MAR-2004 (TREMREL. 26, Last annotation update)
DE Aminoglycoside-O-phosphotransferase type 5.
GN Name=aph;
OS Streptomyces rimosus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1927;
RN [1]
RP SEQUENCE FROM N.A.
RA Akopiants K.E., Danilenko V.N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Akopiants K.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U24442; AAS03856.1; -;
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR002575; APH trans.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF01636; APH; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
SQ TRANSFERASE.
KW SEQUENCE 258 AA; 28490 MW; D2AD3F681470264E CRC64;
Query Match 48.8%; Score 63; DB 2; Length 258;
Best Local Similarity 50.0%; Pred.No. 2.8; Mismatches 7; Indels 0; Gaps 0;
Matches 8; Conservative 1;
QY 3 WSSWALGWRWLRYYGWM 18
Db 56 WQLGPGWAWVRLSCM 71
RESULT 2
Q73VR3 PRELIMINARY; PRT; 312 AA.
ID Q73VR3
AC Q73VR3;
DT 05-JUL-2004 (TREMREL. 27, Created)
DT 05-JUL-2004 (TREMREL. 27, Last sequence update)
DT 05-JUL-2004 (TREMREL. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MAP2949c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;

```

RN  SEQUENCE FROM N.A.
RC  STRAIN=k10;
RA  Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL  Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AE017237; AAS05266.1; -.
DR  InterPro; IPR002123; Acyltransferase.
DR  Pfam; PF01553; Acyltransferase; 1.
DR  SMART; SM00563; PlcC; 1.
KW  Complete proteome; Hypothetical protein.
SQ  SEQUENCE 312 AA; 35001 MW; 577E4AE35A177103 CRC64;

Query Match 47.3%; Score 61; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy  2 VWSSWALGWRWLRYYG 17
Db  101 VWDATVGLQWRRFG 116

RESULT 3
AAS05266 PRELIMINARY; PRT; 312 AA.
AC  AAS05266;
DT  02-MAR-2004 (TrEMBLrel. 27, Created)
DT  02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE  Hypothetical protein.
GN  MAP2949C.
OS  Mycobacterium paratuberculosis.
OC  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC  Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX  NCBI_TaxID=1770;
RN  [1];
RP  SEQUENCE FROM N.A.
RC  STRAIN=k10;
RA  Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL  Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AE017237; AAS05266.1; -.
KW  Hypothetical protein.
SQ  SEQUENCE 312 AA; 35001 MW; 577E4AE35A177103 CRC64;

Query Match 47.3%; Score 61; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy  2 VWSSWALGWRWLRYYG 17
Db  101 VWDATVGLQWRRFG 116

RESULT 4
Q6JN63 PRELIMINARY; PRT; 488 AA.
AC  Q6JN63;
DT  05-JUL-2004 (TrEMBLrel. 27, Created)
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE  Hypothetical protein.
OS  Xanthomonas campestris.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC  Xanthomonadaceae; Xanthomonas.
OX  NCBI_TaxID=339;
RN  [1];
RP  SEQUENCE FROM N.A.
RC  STRAIN=IG8;
RA  Ryan R.P., Ryan D.J., Dowling D.N.;
RL  Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AY359472; AAT39386.1; -.
KW  Hypothetical protein.
SQ  SEQUENCE 488 AA; 52189 MW; 8B16A7998475E512 CRC64;

```

```

Query Match 46.9%; Score 60.5; DB 2; Length 488;
Best Local Similarity 41.4%; Pred. No. 11;
Matches 12; Conservative 1; Mismatches 3; Indels 13; Gaps 2;

Qy  2 VWSSW-----ALGWRWLRYYGW 18
Db  98 VWGSWLLASAPLFVIFNALGW-WLDAFGW 125

RESULT 5
AAT39386 PRELIMINARY; PRT; 488 AA.
AC  AAT39386;
DT  01-JUN-2004 (TrEMBLrel. 27, Created)
DT  01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT  01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE  Hypothetical protein.
OS  Xanthomonas campestris.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC  Xanthomonadaceae; Xanthomonas.
OX  NCBI_TaxID=339;
RN  [1];
RP  SEQUENCE FROM N.A.
RC  STRAIN=IG8;
RA  Ryan R.P., Ryan D.J., Dowling D.N.;
RL  "The Xme-like system of strain Xanthomonas campestris IG8 mediates
RT  metal resistance.";
RL  Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AY359472; AAT39386.1; -.
KW  Hypothetical protein.
SQ  SEQUENCE 488 AA; 52189 MW; 8B16A7998475E512 CRC64;

Query Match 46.9%; Score 60.5; DB 2; Length 488;
Best Local Similarity 41.4%; Pred. No. 11;
Matches 12; Conservative 1; Mismatches 3; Indels 13; Gaps 2;

Qy  2 VWSSW-----ALGWRWLRYYGW 18
Db  98 VWGSWLLASAPLFVIFNALGW-WLDAFGW 125

RESULT 6
Q7UWQ5 PRELIMINARY; PRT; 466 AA.
AC  Q7UWQ5;
DT  01-OCT-2003 (TrEMBLrel. 25, Created)
DT  01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Hypothetical protein.
GN  OrderedLocusNames=RH1870;
OS  Rhodopirellula baltica.
OC  Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC  Planctomycetaceae; Pirellula.
OX  NCBI_TaxID=117;
RN  [1];
RP  SEQUENCE FROM N.A.
RC  STRAIN=1;
RX  MEDLINE=22735913; PubMed=12835416;
RA  Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA  Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA  Schlesner H., Amann R., Reinhardt R.;
RT  "Complete genome sequence of the marine planctomycete Pirellula sp.
RT  strain 1.";
RL  Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR  EMBL; BX294136; CAD72307.1; -.
KW  Complete proteome; Hypothetical protein.
SQ  SEQUENCE 466 AA; 53606 MW; F6FFC58A6FD97A97 CRC64;

Query Match 45.0%; Score 58; DB 2; Length 466;
Best Local Similarity 64.3%; Pred. No. 21;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy  1 QVSSWALGWRWLR 14

```

```

Db      197 QAWSQMSLSKRWLR 210
      | | | | | | | | | |
RESULT 7
QBZ40
ID Q8BZ40 PRELIMINARY; PRT; 164 AA.
AC Q8BZ40;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE Mus musculus adult male bone cDNA, RIKEN full-length enriched library,
clone:9830169C09 product:ALANLY-TRNA SYNTHETASE (EC 6.1.1.7) (ALANINE-
-TRNA LIGASE) (ALARS) homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=21085860; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsumai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,

```

```

RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK036738; BAC29558.1;
DR GO; GO:0004813; F:alanine-tRNA ligase activity; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
KW Aminoacyl-tRNA synthetase; Ligase.
SQ SEQUENCE 164 AA; 18827 MW; D73DC68F0E60F610 CRC64;

Query Match      44.6%; Score 57.5; DB 2; Length 164;
Best Local Similarity 58.8%; Pred. No. 9.1;
Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY      3 WSSNAL-CWRWLRRYGW 18
      | | | | | | | | | |
DB      75 WFWALAGRWLPGMGW 91

RESULT 8
ID Q7U9N8 PRELIMINARY; PRT; 163 AA.
AC Q7U9N8;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Possible 2-amino-4-hydroxy-6-hydroxymethylidihydropteridine
DE pyrophosphokinase (EC 2.7.6.3).
GN Name=folk; OrderedLocusNames=SYNM0216;
OS Synecococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahmasha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synecococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569689; CAE06731.1;
DR GO; GO:0003848; F:2-amino-4-hydroxy-6-hydroxymethylidihydropte. .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009396; F:folic acid and derivative biosynthesis; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF01288; HPPK; 1.
DR TIGRFAMs; TIGR01498; folk; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00794; HPPK; 1.
KW Complete proteome; Kinase; Transferase.
SQ SEQUENCE 163 AA; 17774 MW; 39D2021EA256F92F CRC64;

Query Match      44.2%; Score 57; DB 2; Length 163;
Best Local Similarity 83.3%; Pred. No. 10;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 QWSSWALGNRW 12
      | | | | | | | | | |
DB      42 QWSSQALGFRW 53

RESULT 9
Q933R0
ID Q933R0 PRELIMINARY; PRT; 241 AA.
AC Q933R0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative alkane-1-hydroxylase (Putative alkane-1-monooxygenase)
(Fragment).
GN Name=alkB2;

```

DR	EMBL	A0633617	CAG17609.1.1	-	-	-
DR	EMBL	AJ633618	CAG17610.1.1	-	-	-
DR	EMBL	AJ633619	CAG17611.1.1	-	-	-
DR	EMBL	AJ633620	CAG17612.1.1	-	-	-
DR	InterPro	IPR005804	FA_desat.	-	-	-
DR	Pfam	PF00487	FA_desaturase	1.	-	-
SQ	SEQUENCE	377 AA;	43408 MW;	1F23438DA9F1822E	CR064;	-
Query Match				44.2%;	Score 57;	DB 2; Length 377;
Best Local Similarity				42.9%;	Pred. No. 23;	-

Best Local Similarity 42.9%; Pred. No. 42;
Matches 9; Conservative 2; Mismatches 6; Indels 4; Gaps 1;

Qy 2 VWSSWALG---WRWLRRYGW 18
23 VWSGWILAHETRPDWWGOLGW 43
Db

Q9I3I8

```

RESULT 15
AAR13803
ID AAR13803 PRELIMINARY; PRT; 700 AA.
AC AAR13803;
DT 02-MAR-2004 (T-EMBLrel. 27, Created)
DI 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
DE Alkane hydroxylase.
GN ALKB.
OS Pseudomonas frederiksbergensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
CX NCBI_TaxID=104087;
PN [1]

```

RC STRAIN=ATCC 15692

[illegible]

Best Local Similarity 42.9%; Pred No 33.
44.2%; Score 37; DB 2; Length 377;

06SOM6

Query Match

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:09:56 ; Search time 85.5 Seconds
(without alignments)
83.913 Million cell updates/sec

Title: US-10-066-965A-2

Perfect score: 129

Sequence: 1 QWSSWALCWRWLRYYGWM 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqpl1980s:*

2: Geneseqpl1990s:*

3: Geneseqpl2000s:*

4: Geneseqpl2001s:*

5: Geneseqpl2002s:*

6: Geneseqpl2003as:*

7: Geneseqpl2003bs:*

8: Geneseqpl2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	20	5	AA018008 Intracell
2	117	90.7	20	2	AAR89346 Cdk2-inte
3	117	90.7	20	2	AAW32121 Interacti
4	117	90.7	20	3	AAV49340 Cdk2 inte
5	89.5	69.4	17	5	AA018007 Intracell
6	59	45.7	118	4	ABG24883 Novel hum
7	57	44.2	382	7	ABO83419 Pseudomon
8	57	44.2	1086	7	AAU17861 Novel hum
9	56	43.4	73	4	AAU17861 Novel hum
10	56	43.4	73	7	ADG41241 Human res
11	56	43.4	122	8	ADG22291 Cyanophag
12	55.5	43.0	78	8	ADG22339 Cyanophag
13	55.5	43.0	136	6	ABR56763 Human sec
14	55	42.6	282	2	AAV32836 Woodchuck
15	54.5	42.2	452	5	ABR92178 Herbicida
16	54.5	42.2	574	6	ABJ39104 Molecule
17	54	41.9	68	4	AAW80357 Human hae
18	54	41.9	599	4	ABG18132 Novel hum
19	54	41.9	1725	3	AAW48460 Rat lamin
20	54	41.9	2115	6	ABP70920 Arabidops
21	54	41.9	2150	6	ABP70925 Arabidops
22	54	41.9	2159	6	ABP70924 Maize DEX
23	53	41.1	89	4	ABG06199 Novel hum
24	53	41.1	105	4	ABG11452 Novel hum
25	52.5	40.7	173	7	ABO83613 Pseudomon

ALIGNMENTS

RESULT 1

AA018008

ID AAO18008 standard; peptide; 20 AA.

XX

AC AAO18008;

XX

DT 30-AUG-2002 (first entry)

XX

DE Intracellular target molecule property modulation method aptamer 10M.

XX

KW Intracellular target; cellular component; property modulation;

KW antimicrobial; immunomodulatory; nootropic; neuroprotective; metabolic;

KW neuroleptic; cytostatic; cardiac; infection; immunological disorder;

KW neurological disorder; metabolic disorder; psychiatric disorder;

KW myopathy; cancer; cardiovascular disorder.

XX

OS Unidentified.

XX

PN EP1205191-A1.

XX

PD 15-MAY-2002.

XX

PF 13-NOV-2000; 2000EP-00403156.

XX

PR 13-NOV-2000; 2000EP-00403156.

XX

PA (CNRS) CENT NAT RECH SCI.

XX

PA (MASS-) MASSACHUSETTS GEN HOSPITAL.

XX

PA (MOLE-) MOLECULAR SCI INST.

XX

PI Colas P, Brent R, Cohen BA;

XX

DR WPI; 2002-418829/45.

XX

CC Process for specifically modulating the properties of an intracellular

CC target molecule used for the treatment of various disorders.

XX

XX Example 1; Fig 1; 33pp; English.

XX

CC The present invention relates to a process for specifically modulating

CC the properties of an intracellular target molecule T, and/or of a

CC cellular component C which interacts directly or indirectly in a cell

CC with the target. The process involves the introduction into the cell of a

CC chimeric molecule known as a targeted effector, comprising a recognition

CC moiety capable of recognising T and an effector moiety. The chimeric

CC protein or nucleic acid can be used in the preparation of a medicament

CC for the treatment of microbial infections, immunological disorders,

CC neurological disorders, metabolic disorders, psychiatric disorders,

CC myopathies, genetic disorders, cancer, cardiovascular disorders and
CC dental disorders. The present sequence is a mutant of a known anti-Cdk2
CC aptamer used in the exemplification of the invention

XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 129; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QVSSWALGWRWLRYYGWM 20
|||||
Db 1 QVSSWALGWRWLRYYGWM 20
|||||

RESULT 2

AAR89346
ID AAR89346 standard; peptide; 20 AA.

XX AC AAR89346;

XX DT 10-SEP-1996 (first entry)

XX DE Cdk2-interacting peptide isolated using interaction trap assay.

XX KW Cdk2; cyclin dependent kinase 2; assay; identifying; isolating;
KW cell cycle; interaction; antagonist; conformationally-constrained;
KW agonist; interaction trap; thioredoxin; LexA; two-hybrid system.

XX OS Synthetic.

XX PN WO9602561-A1.

XX PD 01-FEB-1996.

XX PF 20-JUL-1995; 95WO-US009307.

XX PR 20-JUL-1994; 94US-00278082.

XX PA (GEMO) GEN HOSPITAL CORP.

XX PA (GEMO) GENETICS INST INC.

XX PI Brent R, McCoy JM, Jessen TH, Xu C;

XX PS WPI; 1996-105852/11.

XX PT Interaction trap systems using conformationally-constrained proteins -
PT useful for detection of protein interactions and for identification and
PT isolation of interacting proteins.

XX PS Claim 66; Page 60; 73pp; English.

CC AAR89340-R89351 are Cdk-2 interacting peptides identified by an
CC interaction trap assay using conformationally-constrained proteins. The
CC assay comprises providing a host cell (esp. a yeast cell) which contains:
CC (i) a reporter gene operably linked to a DNA-binding protein recognition
CC site, pref. LexA; (ii) a 1st fusion gene expressing Cdk-2 covalently
CC bonded to a LexA binding protein; and (iii) a 2nd fusion gene expressing
CC a conformationally-constrained (pref. with thioredoxin) protein (either
CC expected to interact with Cdk-2 or a random peptide) covalently bonded to
CC a gene-activating moiety, and measuring expression of the reporter gene
CC as a measure of interaction between Cdk2 and the conformationally
CC constrained protein. The same system may be applied to find proteins that
CC interact with any other protein of interest (e.g. Ras) or proteins having
CC agonist or antagonist activity on such interactions

XX SQ Sequence 20 AA;

Query Match 90.7%; Score 117; DB 2; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.9e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 QVSSWALGWRWLRYYGWM 20
|||||

Db 1 QVSSWALGWRWLRYYGWM 20
|||||

RESULT 3

AAW32121
ID AAW32121 standard; peptide; 20 AA.

XX AC AAW32121;

XX DT 22-APR-1998 (first entry)

XX DE Interaction trap system aptamer 7.

XX KW Protein interaction; interaction trap; fusion protein; mimetic;
KW therapeutic; detection; reporter gene.

XX OS Synthetic.

XX PN WO9738127-A1.

XX PD 16-OCT-1997.

XX PF 09-APR-1997; 97WO-US005793.

XX PR 09-APR-1996; 96US-00630052.

XX PA (GEMO) GEN HOSPITAL CORP.

XX PA (GEMO) GENETICS INST INC.

XX PI Brent R, McCoy JM, Jessen TH, Xu CW;

XX PS WPI; 1997-512733/47.

XX PT New trap system for detecting protein interactions - comprises a reporter
PT gene linked to a DNA-binding-protein recognition site and fusion proteins
PT to test for interactions.

XX PS Claim 73; Page 52; 89pp; English.

CC Aptamers AAW32116-W32132 have been isolated from a peptide library and
CC are used in a novel interaction trap method for detecting protein
CC interactions and isolating novel proteins. The method involves a host
CC cell containing a reporter gene operably linked to a DNA-binding-protein
CC (DBP) recognition site, a fusion gene capable of expressing a fusion
CC protein which is able to specifically bind to the DBP recognition site
CC and a second fusion gene which expresses a second fusion protein which is
CC conformationally constrained and bonded to a gene activating moiety.
CC Measuring expression of the reporter gene gives a measure of the
CC interaction between P1 and P2. This method can be used to identify
CC agonists or antagonists for use as therapeutic molecules or for the
CC design of simple molecule mimetics. The method is specifically used to
CC detect an interacting protein in a population of proteins or to identify
CC a candidate interactor. Using conformationally constrained proteins can
CC provide for tertiary structural analysis and can also protect proteins
CC from cellular degradation and/or increase the protein's solubility,
CC and/or otherwise alter the capacity of the candidate interactor to
CC interact

XX SQ Sequence 20 AA;

Query Match 90.7%; Score 117; DB 2; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.9e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 QVSSWALGWRWLRYYGWM 20
|||||

Db 1 QVSSWALGWRWLRYYGWM 20
|||||

RESULT 4

AAV49340

ID AAV49340 standard; peptide; 20 AA.

CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX

SQ Sequence 1086 AA;

Query Match 44.2%; Score 57; DB 7; Length 1086;

Best Local Similarity 62.5%; Pred. No. 75;

Matches 10; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 3 WSSWALGWRWLRYYGW 18

Db |||||
52 WSRWA---WLRNGW 63

RESULT 9

AAU17861

ID AAU17861 standard; protein; 73 AA.

XX

AC AAU17861;

XX

DT 07-NOV-2001 (first entry)

XX

DE Novel human respiratory antigen #177.

XX

KW Human; respiratory antigen; respiratory disorder; throat disorder;

KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;

KW anti allergic; anti asthmatic; anti inflammatory; olfactory;

KW respiratory active.

XX

OS Homo sapiens.

XX

PN WO20015448-A1.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US001333.

XX

PR 31-JAN-2000; 2000US-0179065P.

PR

PR 04-FEB-2000; 2000US-0180628P.

PR

PR 24-FEB-2000; 2000US-0184664P.

PR

PR 02-MAR-2000; 2000US-0186350P.

PR

PR 16-MAR-2000; 2000US-0189874P.

PR

PR 17-MAR-2000; 2000US-0190076P.

PR

PR 18-APR-2000; 2000US-0198123P.

PR

PR 19-MAY-2000; 2000US-0205515P.

PR

PR 07-JUN-2000; 2000US-0209467P.

PR

PR 28-JUN-2000; 2000US-0214886P.

PR

PR 30-JUN-2000; 2000US-0215135P.

PR

PR 07-JUL-2000; 2000US-0216647P.

PR

PR 07-JUL-2000; 2000US-0216880P.

PR

PR 11-JUL-2000; 2000US-0217487P.

PR

PR 11-JUL-2000; 2000US-0217496P.

PR

PR 14-JUL-2000; 2000US-0218290P.

PR

PR 26-JUL-2000; 2000US-0220963P.

PR

PR 26-JUL-2000; 2000US-0220964P.

PR

PR 14-AUG-2000; 2000US-0224518P.

PR

PR 14-AUG-2000; 2000US-0224519P.

PR

PR 14-AUG-2000; 2000US-0225213P.

PR

PR 14-AUG-2000; 2000US-0225214P.

PR

PR 14-AUG-2000; 2000US-0225266P.

PR

PR 14-AUG-2000; 2000US-0225267P.

PR

PR 14-AUG-2000; 2000US-0225268P.

PR

PR 14-AUG-2000; 2000US-0225270P.

PR

PR 14-AUG-2000; 2000US-0225447P.

PR

PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-476224/51.
XX N-PSDB; AAS28045.

PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the respiratory system including respiratory cancers
PT and also for testing and detection e.g. diagnosis.

XX Claim 11; SED ID No 479; 546pp; English.

XX The present invention relates to the isolation of novel human respiratory
CC antigens, and cDNA (AAS27869-AAS28159) and genomic sequences encoding for
CC these polypeptides. The sequences of the invention are useful for
CC preventing, treating and/or prognosing disorders related to the
CC respiratory system including throat disorders (e.g. vocal cord paralysis,
CC tonsillitis, and laryngitis), lung disorders (e.g. pneumonia, allergic
CC disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose
CC disorders and cancers of the respiratory tissues e.g. lung cancer. The
CC polynucleotide sequences of the invention are useful in gene therapy and
CC antisense therapy. AAU17685-AAU17975 represent novel human respiratory
CC antigens. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 73 AA;

Query Match 43.4%; Score 56; DB 4; Length 73;
Best Local Similarity 47.1%; Pred. No. 6.4;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 VWSWALGWRWLRRCYGW 18
DB 16 INSWMLPGQWLRPKSW 32

RESULT 10
ADG41241
ID ADG41241 standard; protein; 73 AA.
XX AC ADG41241;
XX DT 26-FEB-2004 (first entry)
XX DE Human respiratory system associated protein seq id 479.
XX KW antiinflammatory; antiallergic; antiasthmatic; cytostatic; gene therapy;
KW human respiratory system antigen;
KW human respiratory system associated polynucleotide;
KW respiratory system disorder; throat disorder; vocal cord paralysis;
KW tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
KW histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
KW cancer; respiratory tissue cancer; throat cancer; lung cancer;
KW cancer of the nose; gene therapy; chromosome identification; forensic;
KW human respiratory system associated protein; human.
XX OS Homo sapiens.
XX PN US2003215893-A1.
XX PD 20-NOV-2003.
XX PF 07-AUG-2002; 2002US-00212872.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.


```

XX 26-FEB-2004 (first entry)
XX DT
XX PF
XX DE
XX KW
XX genome; cyanophages; 2; 6-diaminopurine; chemotherapy; AIDS.
XX OS
XX Cyanophage S-2L.
XX PN FR2839079-A1.
XX 31-OCT-2003.
XX PD
XX 30-APR-2002; 2002FR-00005424.
XX PF
XX 30-APR-2002; 2002FR-00005424.
XX PR
XX (INSP ) INST PASTEUR.
XX PA (CNRS ) CNRS CENT NAT RECH SCI.
XX PA (GENO-) GENOSCOPE CENT NAT SEQUENCAGE GRP INTERE.
XX PI Marliere P, Kaminski PA, Galisson F, Bouzon M, Pochet S;
XX PI Weissenbach J, Saurin W, Robert C, Vico V;
XX WPI; 2004-045746/05.
XX DR N-PSDB; ADG22255.
XX
XX New genomic sequence for cyanophage S-2L, useful for identifying genes
XX PT for synthesis of 2,6-diaminopurine bases or polynucleotides containing
XX PT them.
XX PS Claim 6; SEQ ID NO 37; 423pp; French.
XX CC The invention relates to the entire genome of cyanophage S-2L, and to the
XX CC protein encoded by it. Genes isolated from the genome of S-2L are useful
XX CC for preparing enzymes for synthesis of D-bases (D = 2,6-diaminopurine),
XX CC particularly D, dUMP and dTTP, or polynucleotides containing these bases,
XX CC polymerases involved in metabolism of D-bases and deoxynucleotide
XX CC analogs, for chemotherapy of AIDS. The genes, and encoded polypeptides,
XX CC can be used for detection and/or identification of D-bases or polynucleotides
XX CC identifying agents that modulate synthesis of D-bases or polynucleotides
XX CC containing them, and fusions of S-2L polypeptides with an antigen can be
XX CC used to raise specific antibodies, useful for detecting S-2L. This
XX CC sequence corresponds to one of the proteins encoded by the cyanophage S-
XX CC 2L genome.
XX SQ Sequence 122 AA;
XX
XX Query Match 43.4%; Score 56; DB 8; Length 122;
XX Best Local Similarity 80.0%; Pred. No. 11;
XX Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 9 GWRWLRRYGW 18
XX |||||
XX 112 GWRWCRGCGW 121
XX
XX RESULT 12
XX ADG22339
XX ID ADG22339 standard; protein; 78 AA.
XX XX
XX AC ADG22339;
XX XX
XX 26-FEB-2004 (first entry)
XX DT
XX DE
XX Cyanophage S-2L encoded protein #84.
XX KW genome; cyanophages; 2; 6-diaminopurine; chemotherapy; AIDS.
XX OS
XX Cyanophage S-2L.
XX PN FR2839079-A1.
XX

```

```

PD 31-OCT-2003.
XX 30-APR-2002; 2002FR-00005424.
XX PF
XX 30-APR-2002; 2002FR-00005424.
XX PR
XX (INSP ) INST PASTEUR.
XX PA (CNRS ) CNRS CENT NAT RECH SCI.
XX PA (GENO-) GENOSCOPE CENT NAT SEQUENCAGE GRP INTERE.
XX PI Marliere P, Kaminski PA, Galisson F, Bouzon M, Pochet S;
XX PI Weissenbach J, Saurin W, Robert C, Vico V;
XX WPI; 2004-045746/05.
XX DR N-PSDB; ADG22255.
XX
XX New genomic sequence for cyanophage S-2L, useful for identifying genes
XX PT for synthesis of 2,6-diaminopurine bases or polynucleotides containing
XX PT them.
XX PS Claim 6; SEQ ID NO 85; 423pp; French.
XX CC The invention relates to the entire genome of cyanophage S-2L, and to the
XX CC protein encoded by it. Genes isolated from the genome of S-2L are useful
XX CC for preparing enzymes for synthesis of D-bases (D = 2,6-diaminopurine),
XX CC particularly D, dUMP and dTTP, or polynucleotides containing these bases,
XX CC polymerases involved in metabolism of D-bases and deoxynucleotide
XX CC analogs, for chemotherapy of AIDS. The genes, and encoded polypeptides,
XX CC can be used for detection and/or identification of D-bases or polynucleotides
XX CC identifying agents that modulate synthesis of D-bases or polynucleotides
XX CC containing them, and fusions of S-2L polypeptides with an antigen can be
XX CC used to raise specific antibodies, useful for detecting S-2L. This
XX CC sequence corresponds to one of the proteins encoded by the cyanophage S-
XX CC 2L genome.
XX SQ Sequence 78 AA;
XX
XX Query Match 43.0%; Score 55.5; DB 8; Length 78;
XX Best Local Similarity 50.0%; Pred. No. 8;
XX Matches 9; Conservative 3; Mismatches 3; Indels 3; Gaps 2;
XX
XX QY 3 WSSWALGWRWLRRYGWGM 20
XX ||:|||||
XX 44 WALWS-AWRWLR--SWGL 58
XX
XX Db
XX
XX RESULT 13
XX ABR56763
XX ID ABR56763 standard; protein; 136 AA.
XX AC
XX ABR56763;
XX XX
XX 30-JUL-2003 (first entry)
XX DT
XX DE
XX Human secreted protein SECP-38 SEQ ID NO:38.
XX KW
XX Human; secreted protein; SECP; cytostatic; antiarteriosclerotic;
XX KW anticonvulsant; nootropic; neuroprotective; cerebroprotective; anti-HIV;
XX KW antiallergic; antiinflammatory; thyromimetic; gene therapy; cancer;
XX KW cell proliferative disorder; atherosclerosis; neurological disorder;
XX KW epilepsy; Huntington's disease; stroke; immune disorder; allergy; AIDS;
XX KW inflammatory disorder; developmental disorder; hypothyroidism;
XX KW Cushing's syndrome; infection.
XX OS
XX Homo sapiens.
XX XX
XX WO2003016506-A2.
XX PN
XX 27-FEB-2003.
XX PD
XX 15-AUG-2002; 2002WO-US027143.
XX PF
XX 17-AUG-2001; 2001US-0313249P.
XX PR

```


PR 24-AUG-2001; 2001US-0314752P.
 PR 07-SEP-2001; 2001US-0317818P.
 PR 07-SEP-2001; 2001US-0317824P.
 PR 21-SEP-2001; 2001US-0324040P.
 PR 24-SEP-2001; 2001US-0324586P.
 PR 02-NOV-2001; 2001US-0343980P.
 PR 28-NOV-2001; 2001US-0334229P.
 PR 13-FEB-2002; 2002US-0357002P.
 PR 06-MAR-2002; 2002US-0362439P.
 PR 19-MAR-2002; 2002US-0366041P.
 PR 30-APR-2002; 2002US-0376988P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Tang TY, Warren BA, Gietzen KJ, Lal PG, Yue H, Honchell CD;
 PI Lehr-Mason PM, Burford N, Xu Y, Baughn MR, Duggan BM, Tran UK;
 PI Lee EA, Forsythe IJ, Richardson TW, Lee S, Thangaveilu K, Yue H;
 PI Emerling BM, Wallia NK, Azimzai Y, Sanjwala B, Hafalia AJA;
 PI Borowsky ML, Nguyen DB, Ison CH, Astromoff A, Ding L, Lee SY;
 PI Becha SD, Ramkumar J, Gandhi AR, Jin P, Fu GK, Swarnakar A;
 XX WPI: 2003-278569/27.
 DR N-PSDB; ACC79063.
 XX
 XX New human secreted proteins (SECP), useful for diagnosing, treating and
 PT preventing diseases or conditions associated with the aberrant SECP
 PT expression e.g. cancer, AIDS, atherosclerosis, Huntington's disease,
 PT stroke, infections.
 XX
 XX Claim 1; Page 217; 286pp; English.
 PS
 XX ACC79026 to ACC79105 encode the human secreted proteins (I) given in
 CC ABR56726 to ABR56805, designated SECP-1 to SECP-80. SECP sequences can
 CC have cytostatic, antiarteriosclerotic, anticonvulsant, antiinflammatory,
 CC neurotropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic and
 CC thymometric activities, and can be used in gene therapy. The SECP
 CC proteins and polynucleotides can be used in diagnosing, treating and
 CC preventing diseases or conditions associated with the decreased
 CC expression or overexpression of SECP, such as cell proliferative (e.g.
 CC cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's
 CC disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and
 CC developmental (e.g. hypothyroidism, Cushing's syndrome) disorders, or
 CC infections. They are also useful in assessing the effects of exogenous
 CC compounds on the expression of nucleic acid and amino acid sequences of
 CC SECP. The SECP or its fragments are useful in screening compounds for
 CC effectiveness as agonist or antagonist of the polypeptides, or in
 CC altering the expression of the target polynucleotide and compounds that
 CC specifically bind to or modulate the activity of the polypeptide
 XX
 SQ Sequence 136 AA;
 Query Match 43.0%; Score 55.5; DB 6; Length 136;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 12; Conservative 1; Mismatches 4; Indels 3; Gaps 2;
 QY 2 VMSWAL--GMRWLRYYGNG 19
 ||| ||| ||| ||| |||
 Db 76 VW-WNPLNPGWGLRWQNG 94
 RESULT 14
 AAY32836
 ID AAY32836 standard; protein; 282 AA.
 XX
 AC AAY32836;
 XX
 DT 17-OCT-2003 (revised)
 DT 29-OCT-1999 (first entry)
 XX
 XX Woodchuck hepatitis b virus HBsAg Pres2-S region protein sequence.
 DE
 XX HBsAg; Pres2-S; recombinant antigen library; disease-related antigen;
 KW multivalent antigenic polypeptide production; infection; allergen;
 XX

KW asthma; autoimmune disease; rheumatoid arthritis; diabetes; therapy;
 KW multiple sclerosis; inflammatory condition; cancer; contraception;
 KW immune response; hepatitis b surface antigen.
 XX
 OS Woodchuck hepatitis B virus.
 XX WO9941383-A1.
 PN
 XX
 PD 19-AUG-1999.
 XX
 XX 10-FEB-1999; 99WO-US002944.
 XX
 PR 11-FEB-1998; 98US-00021769.
 PR 11-FEB-1998; 98US-0074294P.
 PR 23-OCT-1998; 98US-0105509P.
 XX (MAXY-) MAXYGEN INC.
 PA
 XX Punnonen J, Bass SH, Whalen RG, Howard R, Stemmer WPC;
 PI WPI: 1999-518452/43.
 DR N-PSDB; AAZ10971.
 XX
 PT Recombinant multivalent antigenic polypeptide produced by recombining
 PT nucleic acid sequences and screening, used in vaccines against e.g.
 PT infections and cancer.
 XX
 PS Example 14; Fig 17; 153pp; English.
 XX
 CC This sequence is the woodchuck hepatitis virus surface antigen Pres2-S
 CC region. This sequence was used to create a recombinant antigen library.
 CC The library comprises recombinant nucleic acids encoding antigenic
 CC polypeptides and is produced by recombination of at least two forms of
 CC nucleic acid, differing by at least two nucleotides, encoding a disease-
 CC related antigenic polypeptide. The library can be used to produce a
 CC recombinant multivalent antigenic polypeptides of the invention, that
 CC contains at least two antigenic determinants (AD) from different
 CC polypeptides. The multivalent antigenic polypeptides are used in vaccines
 CC to induce a protective or therapeutic response to a wide variety of
 CC infectious agents (bacteria, viruses, parasites, including Plasmodium
 CC falciparum); allergens; asthma; autoimmune disease (e.g. rheumatoid
 CC arthritis, diabetes, multiple sclerosis); other inflammatory conditions
 CC and cancer, also, where directed against sperm antigens, they can be used
 CC for contraception. The multivalent peptides can be evolved to induce an
 CC optimised immune response against a wide variety of antigens.
 CC particularly a broad spectrum response to many different strains of a
 CC pathogen, including strains that are likely to appear in the future.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 282 AA;
 Query Match 42.6%; Score 55; DB 2; Length 282;
 Best Local Similarity 57.9%; Pred. No. 34;
 Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 2;
 QY 4 SSWALG---WRW-LRRYCW 18
 ||||| ||| ||| ||| |||
 Db 210 SSWALGNVLEWALRFSW 228
 RESULT 15
 ABB92178
 ID ABB92178 standard; protein; 452 AA.
 XX
 AC ABB92178;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 XX Herbicidally active polypeptide SEQ ID NO 1389.
 DE
 XX Herbicidal; plant; agriculture; herbicide.
 KW
 XX Arabidopsis thaliana.
 OS

```
XX WO200210210-A2.
XX PN
XX XX
XX PD 07-FEB-2002.
XX XX
XX PF 28-AUG-2001; 2001WO-EP009892.
XX XX
XX PR 28-AUG-2001; 2001WO-EP009892.
XX XX
XX PA (FARB ) BAYER AG.
XX XX
XX EI Tietjen K, Weidler M;
XX DR WPI; 2002-269010/31.
XX XX
XX PT Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.
XX PS
XX SQ Claim 5; SEQ ID NO 1389; 261pp + Sequence Listing; English.
XX CC The invention relates to identifying target proteins (ABB90790-ABB94016)
XX for herbicidally active compounds, comprising aligning and comparing
XX nucleic acid or amino acid sequences from plant with nucleic acid or
XX amino acid sequences from non-plant organisms using suitable search
XX parameters, where plant sequences having an E-value greater by a factor
XX of 3 than the E-value of most similar non-plant sequences are selected.
XX The polypeptides or nucleic acids encoding them are useful for
XX identifying modulators. The identified modulators are useful as
XX herbicides
XX SQ Sequence 452 AA;
XX Query Match 42.2%; Score 54.5; DB 5; Length 452;
XX Best Local Similarity 56.2%; Pred. No. 64;
XX Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;
OY 6 WALGWRWLR-YGWGM 20
Db 77 WTLGWSWAKREVINGM 92
```

Search completed: October 26, 2004, 06:45:09
Job time : 86.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:45:25 ; Search time 64.25 Seconds
(without alignments)
100.781 Million cell updates/sec

Title: US-10-066-965A-2

Perfect score: 129

Sequence: 1 QVWSSWALGWLRRRYGWM 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	129	100.0	20	14	US-10-066-965A-2
2	129	100.0	20	14	US-10-066-965A-8
3	117	90.7	20	14	US-10-162-538-12
4	117	90.7	20	14	US-10-066-965A-1
5	117	90.7	20	14	US-10-066-965A-7
6	65	50.4	135	15	US-10-424-599-149236
7	56	43.4	73	9	US-09-764-860-479
8	56	43.4	73	14	US-10-074-095-479
9	56	43.4	73	14	US-10-212-872-479
10	55	42.6	282	9	US-09-247-890-16
11	55	42.6	282	15	US-10-383-317-16
12	54	41.9	68	9	US-09-796-692-721
13	54	41.9	68	14	US-10-040-862-721

14	54	41.9	68	15	US-10-057-475B-721	Sequence 721, App
15	54	41.9	68	15	US-10-154-884B-721	Sequence 721, App
16	54	41.9	68	15	US-10-764-324-721	Sequence 721, App
17	54	41.9	85	15	US-10-424-599-280831	Sequence 280831,
18	54	41.9	1725	15	US-10-603-725-10	Sequence 10, Appl
19	54	41.9	2094	16	US-10-437-963-187850	Sequence 187850,
20	54	41.9	2115	14	US-10-208-948-4	Sequence 4, Appli
21	54	41.9	2150	14	US-10-208-948-27	Sequence 27, Appli
22	54	41.9	2157	17	US-10-696-616-16	Sequence 16, Appl
23	54	41.9	2159	14	US-10-208-948-24	Sequence 24, Appl
24	52.5	40.7	123	16	US-10-437-963-180497	Sequence 180497,
25	52	40.3	65	16	US-10-437-963-136125	Sequence 136125,
26	52	40.3	105	14	US-10-029-386-32277	Sequence 32277, A
27	52	40.3	152	15	US-10-425-114-54667	Sequence 54667, A
28	52	40.3	191	16	US-10-437-963-150166	Sequence 150166,
29	52	40.3	537	16	US-10-437-963-156540	Sequence 156540,
30	51.5	39.9	242	9	US-09-764-853-609	Sequence 609, App
31	51.5	39.9	242	9	US-09-764-898-199	Sequence 199, App
32	51.5	39.9	410	14	US-10-369-493-4332	Sequence 4332, Ap
33	51.5	39.9	410	14	US-10-369-493-7088	Sequence 7088, Ap
34	51	39.5	239	15	US-10-437-963-191253	Sequence 191253,
35	51	39.5	261	15	US-10-282-122A-50852	Sequence 50852, A
36	51	39.5	346	9	US-09-812-862-2	Sequence 2, Appli
37	51	39.5	513	16	US-10-437-963-142773	Sequence 142773,
38	51	39.5	539	16	US-10-437-963-142769	Sequence 142769,
39	50.5	39.1	104	16	US-10-437-963-106350	Sequence 106350,
40	50	38.8	45	9	US-09-864-761-41282	Sequence 41282, A
41	50	38.8	118	16	US-10-437-963-202613	Sequence 202613,
42	50	38.8	157	16	US-10-437-963-169562	Sequence 169562,
43	50	38.8	219	16	US-10-767-701-52070	Sequence 52070, A
44	50	38.8	251	15	US-10-424-599-203683	Sequence 203683,
45	50	38.8	2160	16	US-10-437-963-187848	Sequence 187848,

ALIGNMENTS

RESULT 1

US-10-066-965A-2

; Sequence 2, Application US/10066965A

; Publication No. US20030143626A1

; GENERAL INFORMATION:

; APPLICANT: COLAS, PIERRE

; APPLICANT: BRENT, ROGER

; APPLICANT: COHEN, BARAK A.

; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS

; FILE REFERENCE: EGYPT 3.0-015

; CURRENT APPLICATION NUMBER: US/10/066,965A

; CURRENT FILING DATE: 2002-12-09

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; OTHER INFORMATION: peptide

US-10-066-965A-2

Query Match 100.0%; Score 129; DB 14; Length 20;

Best Local Similarity 100.0%; Pred. No. 8e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVWSSWALGWLRRRYGWM 20

Db 1 QVWSSWALGWLRRRYGWM 20

RESULT 2

US-10-066-965A-8

; Sequence 8, Application US/10066965A

; Publication No. US20030143626A1

```
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-8

Query Match          100.0%; Score 129; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QVSSSWALGWRWLRYYGWM 20
Db      1 QVSSSWALGWRWLRYYGWM 20
|||||

RESULT 3
US-10-162-538-12
; Sequence 12, Application US/10162538
; Publication No. US20030113749A1
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: McCoy, John M.
; APPLICANT: Jessen, Timm H.
; APPLICANT: Xu, Channing Wilson
; TITLE OF INVENTION: INTERACTION TRAP SYSTEMS FOR DETECTING
; PROTEIN
; INTERACTIONS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/162,538
; FILING DATE: 04-Jun-2002
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,052
; FILING DATE: <unknown>
; APPLICATION NUMBER: 08/504,538
; FILING DATE: July 20, 1995
; APPLICATION NUMBER: 08/278,082
; FILING DATE: July 20, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Karen F. Lech
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/311001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 12:
;
; Query Match          90.7%; Score 117; DB 14; Length 20;
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS: No. US20030113749A1 Relevant
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-162-538-12

Query Match          90.7%; Score 117; DB 14; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.7e-06;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 QVSSSWALGWRWLRYYGWM 20
Db      1 QVSSSWALGWRWLRYYGWM 20
|||||

RESULT 4
US-10-066-965A-1
; Sequence 1, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-1

Query Match          90.7%; Score 117; DB 14; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.7e-06;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 QVSSSWALGWRWLRYYGWM 20
Db      1 QVSSSWALGWRWLRYYGWM 20
|||||

RESULT 5
US-10-066-965A-7
; Sequence 7, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-7

Query Match          90.7%; Score 117; DB 14; Length 20;
```

Best Local Similarity 90.0%; Pred. No. 1.7e-06;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QVSSWALGWRWLRYYGWM 20
|||||
Db 1 QVSSWALGWRWLRYYGWM 20

RESULT 6

US-10-424-599-149236
; Sequence 149236, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149236
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105784C.1.pep
US-10-424-599-149236

Query Match 50.4%; Score 65; DB 15; Length 135;

Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 WSSWALGWRWLRYYGWM 20
|||
Db 118 WRKWCGRWGRWAGV 135

RESULT 7

US-09-764-860-479
; Sequence 479, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 479
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (70)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-860-479

Query Match 43.4%; Score 56; DB 9; Length 73;

Best Local Similarity 47.1%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 VWSSWALGWRWLRYYGWM 18
:|:|:|:|:|:|:|:|:|
Db 16 IWSMPLGQWLRXPXSW 32

RESULT 8

US-10-074-095-479
; Sequence 479, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; Prior APPLICATION NUMBER: 09/764,860
; Prior FILING DATE: 2001-01-17
; Prior APPLICATION NUMBER: 60/179,065
; Prior FILING DATE: 2000-01-31
; Prior APPLICATION NUMBER: 60/180,628
; Prior FILING DATE: 2000-02-04
; Prior APPLICATION NUMBER: 60/214,886
; Prior FILING DATE: 2000-06-28
; Prior APPLICATION NUMBER: 60/217,487
; Prior FILING DATE: 2000-07-11
; Prior APPLICATION NUMBER: 60/225,758
; Prior FILING DATE: 2000-08-14
; Prior APPLICATION NUMBER: 60/220,963
; Prior FILING DATE: 2000-07-26
; Prior APPLICATION NUMBER: 60/217,496
; Prior FILING DATE: 2000-07-11
; Prior APPLICATION NUMBER: 60/225,447
; Prior FILING DATE: 2000-08-14
; Prior APPLICATION NUMBER: 60/218,290
; Prior FILING DATE: 2000-07-14
; Prior APPLICATION NUMBER: 60/225,757
; Prior FILING DATE: 2000-08-14
; Prior APPLICATION NUMBER: 60/226,868
; Prior FILING DATE: 2000-08-22
; Prior APPLICATION NUMBER: 60/216,647
; Prior FILING DATE: 2000-07-07
; Prior APPLICATION NUMBER: 60/225,267
; Prior FILING DATE: 2000-08-14
; Prior APPLICATION NUMBER: 60/216,880
; Prior FILING DATE: 2000-07-07
; Prior APPLICATION NUMBER: 60/225,270
; Prior FILING DATE: 2000-08-14
; Prior APPLICATION NUMBER: 60/251,869
; Prior FILING DATE: 2000-12-08
; Prior APPLICATION NUMBER: 60/235,834
; Prior FILING DATE: 2000-09-27
; Prior APPLICATION NUMBER: 60/234,274
; Prior FILING DATE: 2000-09-21
; Prior APPLICATION NUMBER: 60/234,223
; Prior FILING DATE: 2000-09-21
; Prior APPLICATION NUMBER: 60/228,924
; Prior FILING DATE: 2000-08-30
; Prior APPLICATION NUMBER: 60/224,518
; Prior FILING DATE: 2000-08-14
; Prior APPLICATION NUMBER: 60/236,369
; Prior FILING DATE: 2000-09-29
; Prior APPLICATION NUMBER: 60/224,519
; Prior FILING DATE: 2000-08-14
; Prior APPLICATION NUMBER: 60/220,964
; Prior FILING DATE: 2000-07-26
; Prior APPLICATION NUMBER: 60/241,809
; Prior FILING DATE: 2000-10-20
; Prior APPLICATION NUMBER: 60/249,299
; Prior FILING DATE: 2000-11-17
; Prior APPLICATION NUMBER: 60/236,327
; Prior FILING DATE: 2000-09-29
; Prior APPLICATION NUMBER: 60/241,785
; Prior FILING DATE: 2000-10-20
; Prior APPLICATION NUMBER: 60/244,617
; Prior FILING DATE: 2000-11-01
; Prior APPLICATION NUMBER: 60/225,268

; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17

; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match 43.4%; Score 56; DB 14; Length 73;
Best Local Similarity 47.1%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 VWSWALGWEWLRYYGW 18
Db 16 IWSWLPGEQWLRPSW 32

RESULT 9

US-10-212-872-479
; Sequence 479, Application US/10212872
; Publication No. US20030215893A1
; GENERAL INFORMATION:

```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008AC2
; CURRENT APPLICATION NUMBER: US/10/212,872
; CURRENT FILING DATE: 2002-08-07
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 479
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (30)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (70)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-872-479

```

```

Query Match          43.4%; Score 56; DB 14; Length 73;
Best Local Similarity 47.1%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 2 VWSWALGWRWLRYYGW 18
   :|||:|:|:|:|:|:|:|
Db 16 IWSMPLGQWLRPXSW 32

```

```

RESULT 10
US-09-247-890-16
; Sequence 16, Application US/09247890
; Publication No. US20020198162A1
; GENERAL INFORMATION:
; APPLICANT: Punnonen, Juha
; APPLICANT: Bass, Steven H.
; APPLICANT: Whalen, Robert Gerald
; APPLICANT: Howard, Russell
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Maxygen, Inc.
; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/09/247,890
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: US 60/074,294
; EARLIER FILING DATE: 1998-02-11
; EARLIER APPLICATION NUMBER: US 60/105,509
; EARLIER FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Woodchuck hepatitis B virus
US-09-247-890-16

```

```

Query Match          42.6%; Score 55; DB 9; Length 282;
Best Local Similarity 57.9%; Pred. No. 11e+02;
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 2;

```

```

QY 4 SSWALG---RW-LRRYGW 18
   |||||:|:|:|:|:|:|
Db 210 SSWALGNVLEWALARFSW 228

```

```

RESULT 11
US-10-383-317-16
; Sequence 16, Application US/10383317
; Publication No. US20040001849A1
; GENERAL INFORMATION:
; APPLICANT: Punnonen, Juha

```

```

; APPLICANT: Bass, Steven H.
; APPLICANT: Whalen, Robert Gerald
; APPLICANT: Howard, Russell
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Maxygen, Inc.
; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/10/383,317
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: US/09/247,890
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/074,294
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-11
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/105,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Woodchuck hepatitis B virus
US-10-383-317-16

```

```

Query Match          42.6%; Score 55; DB 15; Length 282;
Best Local Similarity 57.9%; Pred. No. 1.1e+02;
Matches 11; Conservative 1; Mismatches 3; Indels 4; Gaps 2;

```

```

QY 4 SSWALG---RW-LRRYGW 18
   |||||:|:|:|:|:|:|
Db 210 SSWALGNVLEWALARFSW 228

```

```

RESULT 12
US-09-796-692-721
; Sequence 721, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 721
; LENGTH: 68

```

```
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: variant
/ LOCATION: (1)...(68)
/ OTHER INFORMATION: Xaa = Any amino acid
US-09-796-692-721
```

```
Query Match          41.9%; Score 54; DB 9; Length 68;
Best Local Similarity 27.0%; Pred. No. 43;
Matches 10; Conservative 4; Mismatches 3; Indels 20; Gaps 1;
```

```
Qy      3 WSSWA-----LQWRWLLRRYWG 19
      :|||
      :|||
Db      8 WAPWAGRCVTRFGLCREKIGSYGDLGWRWWRQHPAG 44
```

RESULT 13

```
US-10-040-862-721
; Sequence 721, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 721
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(68)
; OTHER INFORMATION: Xaa = Any amino acid
US-10-040-862-721
```

```
Query Match          41.9%; Score 54; DB 14; Length 68;
Best Local Similarity 27.0%; Pred. No. 43;
Matches 10; Conservative 4; Mismatches 3; Indels 20; Gaps 1;
```

```
Qy      3 WSSWA-----LQWRWLLRRYWG 19
      :|||
      :|||
Db      8 WAPWAGRCVTRFGLCREKIGSYGDLGWRWWRQHPAG 44
```

RESULT 14

```
US-10-057-475B-721
; Sequence 721, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 721
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(68)
; OTHER INFORMATION: Xaa = Any amino acid
US-10-057-475B-721
```

```
Query Match          41.9%; Score 54; DB 15; Length 68;
Best Local Similarity 27.0%; Pred. No. 43;
Matches 10; Conservative 4; Mismatches 3; Indels 20; Gaps 1;
```

```
Qy      3 WSSWA-----LQWRWLLRRYWG 19
      :|||
      :|||
Db      8 WAPWAGRCVTRFGLCREKIGSYGDLGWRWWRQHPAG 44
```

RESULT 15

```
US-10-154-884B-721
; Sequence 721, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
```



```

; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 721
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(68)
; OTHER INFORMATION: Xaa = Any amino acid
US-10-154-884B-721

```

Query Match 41.9%; Score 54; DB 15; Length 68;
 Best Local Similarity 27.0%; Pred. NO. 43;
 Matches 10; Conservative 4; Mismatches 3; Indels 20; Gaps 1;

```
QY 3 WSWA-----LGRWLRRYCGW 19
    |:-||
    |||:|:|:|
Db 8 WAPWANGCVTRFGLCREKIGSYGDLGWWRMOHPAG 44
```

Search completed: October 26, 2004, 07:10:18
Job time : 64.25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:20:00 ; Search time 22 Seconds
(without alignments)
60.289 Million cell updates/sec

Title: US-10-066-965A-2

Perfect score: 129

Sequence: 1 QVWSSWALGWRLLRYGWM 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	90.7	20	US-08-504-538A-12	Sequence 12, Appl
2	117	90.7	20	US-08-630-052-12	Sequence 12, Appl
3	117	90.7	20	PCT-US95-09307-12	Sequence 12, Appl
4	57	44.2	382	US-09-252-991A-32165	Sequence 32165, A
5	57	44.2	1086	US-09-252-991A-25051	Sequence 25051, A
6	55	42.6	282	US-09-247-890-16	Sequence 16, Appl
7	55	42.6	282	US-09-724-969-16	Sequence 16, Appl
8	55	42.6	282	US-09-724-852-16	Sequence 16, Appl
9	54	41.9	550	US-08-816-155B-44	Sequence 44, Appl
10	54	41.9	550	US-09-079-587-44	Sequence 44, Appl
11	54	41.9	1725	US-09-560-385A-10	Sequence 10, Appl
12	52.5	40.7	173	US-09-252-991A-32359	Sequence 32359, A
13	51.5	39.9	115	US-09-489-039A-9146	Sequence 9146, Ap
14	51	39.5	346	PCT-US96-10602-2	Sequence 2, Appl
15	51	39.5	480	US-09-252-991A-31470	Sequence 31470, A
16	50	38.8	154	US-09-193-104-26	Sequence 26, Appl
17	50	38.8	446	US-09-252-991A-21791	Sequence 21791, A
18	49.5	38.4	307	US-09-489-039A-12110	Sequence 12110, A
19	49.5	38.4	396	US-09-252-991A-31085	Sequence 31085, A
20	49	38.0	135	US-09-270-767-31648	Sequence 31648, A
21	49	38.0	135	US-09-270-767-46865	Sequence 46865, A
22	49	38.0	138	US-09-252-991A-32834	Sequence 32834, A
23	49	38.0	266	US-09-328-352-6190	Sequence 6190, Ap
24	48	37.2	133	US-09-252-991A-21161	Sequence 21161, A
25	48	37.2	173	US-09-252-991A-32831	Sequence 32831, A
26	47.5	36.8	153	US-09-252-991A-18571	Sequence 18571, A
27	47.5	36.8	201	US-09-252-991A-27351	Sequence 27351, A

28 47 36.4 24 3 US-09-082-279B-1199 Sequence 1199, Ap

29 47 36.4 24 3 US-09-315-304B-1199 Sequence 1199, Ap

30 47 36.4 24 4 US-09-834-784-1199 Sequence 1199, Ap

31 47 36.4 24 4 US-09-515-965A-1199 Sequence 1199, Ap

32 47 36.4 24 4 US-09-350-641C-1199 Sequence 1199, Ap

33 47 36.4 24 4 US-09-350-841A-1199 Sequence 1199, Ap

34 47 36.4 128 4 US-09-252-991A-27977 Sequence 27977, A

35 47 36.4 179 4 US-09-252-991A-33085 Sequence 33085, A

36 47 36.4 174 4 US-09-252-991A-21165 Sequence 21165, A

37 47 36.4 182 4 US-09-252-991A-17896 Sequence 17896, A

38 47 36.4 185 4 US-09-252-991A-20977 Sequence 20977, A

39 47 36.4 191 4 US-09-252-991A-18732 Sequence 18732, A

40 47 36.4 255 4 US-09-252-991A-25838 Sequence 25838, A

41 47 36.4 260 4 US-09-252-991A-22986 Sequence 22986, A

42 47 36.4 275 4 US-09-252-991A-16825 Sequence 16825, A

43 47 36.4 345 3 US-09-291-922-14 Sequence 14, Appl

44 47 36.4 367 2 US-08-655-704B-17 Sequence 17, Appl

45 47 36.4 367 3 US-09-107-755-17 Sequence 17, Appl

ALIGNMENTS

RESULT 1

US-08-504-538A-12

; Sequence 12, Application US/08504538A

; Patent No. 6004746

; GENERAL INFORMATION:

; APPLICANT: Brent, Roger

; APPLICANT: McCoy, John M.

; APPLICANT: Jessen, Timm H.

; TITLE OF INVENTION: INTERACTION TRAP SYSTEMS FOR DETECTING

; TITLE OF INVENTION: PROTEIN INTERACTIONS

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Clark & Elbing LLP

; STREET: 176 Federal Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02110-2214

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/504,538A

; FILING DATE: 07/20/95

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/278,082

; FILING DATE: 07/20/94

; ATTORNEY/AGENT INFORMATION:

; NAME: Paul T. Clark

; REGISTRATION NUMBER: 30,162

; REFERENCE/DOCKET NUMBER: 00786/259001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 428-0200

; TELEFAX: (617) 428-7045

; TELEX:

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 20

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-504-538A-12

Query Match 90.7%; Score 117; DB 3; Length 20;

Best Local Similarity 90.0%; Pred. No. 5e-08;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

US-08-816-155B-44
; Sequence 44, Application US/08816155B
; Patent No. 5990091
; GENERAL INFORMATION:
; APPLICANT: TARTAGLIA, JAMES
; APPLICANT: COX, WILLIAM I.
; APPLICANT: GETTIG, RUSSELL R.
; APPLICANT: MARTINEZ, HECTOR
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: PINCUS, STEVEN E.
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 FIFTH AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,155B
; FILING DATE: 12-MAR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, THOMAS J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454310-2990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-816-155B-44

Query Match 41.9%; Score 54; DB 2; Length 550;
Best Local Similarity 53.8%; Pred. No. 41;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 SSWALGWRWLRRY 16
: | ||||| :
Db 524 TGRWGWRLHGY 536

RESULT 10
US-09-079-587-44
; Sequence 44, Application US/09079587
; Patent No. 6130066
; GENERAL INFORMATION:
; APPLICANT: TARTAGLIA, JAMES
; APPLICANT: COX, WILLIAM I.
; APPLICANT: GETTIG, RUSSELL R.
; APPLICANT: MARTINEZ, HECTOR
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: PINCUS, STEVEN E.
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 FIFTH AVENUE
; CITY: NEW YORK
; STATE: NEW YORK

COUNTRY: USA
ZIP: 10151
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,587
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/816,155
FILING DATE: 12-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: KOWALSKI, THOMAS J.
REGISTRATION NUMBER: 32,147
REFERENCE/DOCKET NUMBER: 454310-2990
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-588-0800
TELEFAX: 212-588-0500
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 550 amino acids
TYPE: amino acid
STRANDEDNESS: n/a
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-09-079-587-44

Query Match 41.9%; Score 54; DB 3; Length 550;
Best Local Similarity 53.8%; Pred. No. 41;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 SSWALGWRWLRRY 16
: | ||||| :
Db 524 TGRWGWRLHGY 536

RESULT 11
US-09-560-385A-10
; Sequence 10, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boutand, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C
; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1725
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-560-385A-10

Query Match 41.9%; Score 54; DB 4; Length 1725;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 WSSWALGWRWL 13
: | ||||| :
Db 7 WSAWCTGWLWI 17

RESULT 12
US-09-252-991A-32359
; Sequence 32359, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32359
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32359

Query Match 40.7%; Score 52.5; DB 4; Length 173;
Best Local Similarity 58.8%; Pred. No. 21;
Matches 10; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 3 WSSWALGNWLRYYGNG 19
||| :||| :||
Db 24 WSSRGSDWRRRC-WG 39

RESULT 13
US-09-489-039A-9146
; Sequence 9146, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9146
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9146

Query Match 39.9%; Score 51.5; DB 4; Length 115;
Best Local Similarity 71.4%; Pred. No. 18;
Matches 10; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 6 WALGWRWLRYYGNG 19
||| :||| :||
Db 30 WALG-QWLRKRGFG 42

RESULT 14
PCT-US96-10602-2
; Sequence 2, Application PC/TUS9610602
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/10602
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017,814
; FILING DATE: 20-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/282001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 346 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-10602-2

Query Match 39.5%; Score 51; DB 5; Length 346;
Best Local Similarity 57.9%; Pred. No. 60;
Matches 11; Conservative 0; Mismatches 4; Indels 4; Gaps 2;

QY 4 SSWALG---WRW-LRRYCW 18
||| :||| :||
Db 274 SSWALGNLWELARLSW 292

RESULT 15
US-09-252-991A-31470
; Sequence 31470, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31470
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31470

Query Match 39.5%; Score 51; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LGWRWLR 15
||| :||| :||
Db 255 LGWRWLR 262

Search completed: October 26, 2004, 06:47:56
Job time : 22 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:19:49 ; Search time 17 Seconds
(without alignments)
113.196 Million cell updates/sec

Title: US-10-066-965A-3

Perfect score: 118

Sequence: 1 PRGAPMMRWVCQMLETFML 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	41.5	429	G84176	glutamate dehydrog
2	49	41.5	545	S23582	alpha-galactosidas
3	48	40.7	359	AH3465	oxidoreductase (EC
4	48	40.7	435	S18609	glutamate dehydrog
5	47	39.8	382	C83337	hypothetical prote
6	46	39.0	476	E72079	phospholipase D ho
7	46	39.0	476	A86545	phospholipase D ty
8	46	39.0	777	S65543	3',5'-cyclic-nucle
9	45	38.1	660	H70798	probable cation-tr
10	45	38.1	1123	S20497	phytochrome A - po
11	45	38.1	1191	AE2501	WD-repeat protein
12	44	37.3	194	AE3633	farnesyl cysteine
13	44	37.3	264	A82100	phosphonate ABC tr
14	44	37.3	350	T12260	caffeoyl-CoA O-net
15	44	37.3	366	S40146	catechol O-methyl
16	44	37.3	370	B95420	probable sideropho
17	44	37.3	400	T25605	hypothetical prote
18	44	37.3	463	T02481	probable protein k
19	44	37.3	499	E86206	hypothetical prote
20	44	37.3	500	T16630	hypothetical prote
21	44	37.3	541	T40518	major facilitator
22	44	37.3	591	S04401	spherulene monoox
23	44	37.3	972	T77454	cation-transportin
24	43.5	36.9	277	T46209	hypothetical prote
25	43.5	36.9	570	T38489	helicase - fission
26	43.5	36.9	614	D72373	hypothetical prote
27	43	36.4	106	T44499	hypothetical prote
28	43	36.4	172	E70057	hypothetical prote
29	43	36.4	368	G86412	F28N24.24 protein

30	43	36.4	430	2	E90880	probable enzymes [
31	43	36.4	430	2	D85738	probable enzymes y
32	43	36.4	430	2	F64892	probable membrane
33	43	36.4	439	2	T49739	related to cyclohe
34	43	36.4	661	2	F70751	hypothetical prote
35	42.5	36.0	208	2	E86014	probable enzyme yh
36	42.5	36.0	208	2	E91168	probable enzyme [i
37	42.5	36.0	208	2	S47687	hypothetical prote
38	42	35.6	184	2	S65818	RNA-directed DNA p
39	42	35.6	203	2	B90266	conserved hypothet
40	42	35.6	230	2	T16479	hypothetical prote
41	42	35.6	361	2	T48538	hypothetical prote
42	42	35.6	364	2	I39525	fructose-bisphosph
43	42	35.6	364	2	I39556	fructose-bisphosph
44	42	35.6	412	2	A70839	probable membrane
45	42	35.6	420	2	T36532	

ALIGNMENTS

RESULT 1

G84176
glutamate dehydrogenase [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: G84176
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Lia
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: G84176
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-429 <STO>
A:Cross-references: UNIPROT:Q9HSM4; GB:AE004437; NID:gi0579809; PIDN:AAG18779.1; GSPDB:G
C:Genetics:
A:Gene: gdhB
C:Superfamily: glutamate dehydrogenase (NAD(P)+)

Query Match 41.5%; Score 49; DB 2; Length 429;
Best Local Similarity 41.2%; Pred. No. 9;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 GAPMMRWVCQMLETFM 19
DB 99 GLSMWMTWKCAVMDLPF 115

RESULT 2

S23582
alpha-galactosidase (EC 3.2.1.22) precursor - Aspergillus niger
C:Species: Aspergillus niger
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S23582; S25653; S18851
R:den Heider, I.F.; Rosell, A.M.M.; van Zuijlen, C.M.; Punt, P.J.; van den Hondel, C.A.M.
Mol. Gen. Genet. 233, 404-410, 1992
A:Title: Cloning and expression of a member of the Aspergillus niger gene family encoding
A:Reference number: S23582; MUID:92318893; PMID:1320186
A:Accession: S23582
A:Molecule type: DNA
A:Residues: 1-545 <DEN1>
A:Cross-references: UNIPROT:P28351; EMBL:X63348; NID:g2327; PIDN:CAA44950.1; PID:g2328
A:Accession: S25653
A:Molecule type: protein
A:Residues: 32-59 <DEN2>
A:Note: it is uncertain whether Met-1 or Met-9 is the initiator
C:Genetics:
A:Gene: agla
C:Superfamily: alpha-galactosidase

C;Keywords: glycoprotein; glycosidase; hydrolase
 F;1-31/Domain: (or 9-31) signal sequence #status predicted <SIG>
 F;32-545/Product: alpha-galactosidase #status experimental <WAT>
 F;57,95,101,131,211,363,444/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.5%; Score 49; DB 2; Length 545;
 Best Local Similarity 42.1%; Pred. No. 11;
 Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 PRGAPMNRWVCQMLETMF 19
 ||| ||| : ||| : ||| :
 Db 43 PMGFNNWARMCDLNETLF 61
 ||| ||| : ||| : ||| : ||| :

RESULT 3
 AH3465
 oxidoreductase (EC 1.1.1.1-) [imported] - Brucella melitensis (strain 16M)
 C;Species: Brucella melitensis
 C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 C;Accession: AH3465
 R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, S.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A;Reference number: AD3252; PMID:11756688
 A;Accession: AH3465
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-359 <KUR>
 A;Cross-references: UNIPROT:Q8YFL7; UNIPROT:Q8G2S3; GB:AE008917; PIDN:AAL52891.1; PID:gl
 A;Experimental source: strain 16M
 C;Genetics:
 A;Gene: BMEI1710
 A;Map position: 1
 C;Keywords: oxidoreductase

Query Match 40.7%; Score 48; DB 2; Length 359;
 Best Local Similarity 45.8%; Pred. No. 11;
 Matches 11; Conservative 2; Mismatches 3; Indels 8; Gaps 2;

QY 1 PRGAP--MMWR-----WVCQMLE 16
 ||| ||| : ||| : ||| : ||| :
 Db 276 PRGEPDELWIRPKGGDDWVCVPLE 299
 ||| ||| : ||| : ||| : ||| :

RESULT 4
 S18609
 glutamate dehydrogenase (NADP) (EC 1.4.1.4) - Halobacterium salinarum
 N;Alternate names: glutamic dehydrogenase; NADP-specific glutamate dehydrogenase
 C;Species: Halobacterium salinarum
 C;Date: 13-Jan-1995 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004
 C;Accession: S18609
 R;Benachhou, N.; Baldacci, G.
 Mol. Gen. Genet. 230, 345-352, 1991
 A;Title: The gene for a halophilic glutamate dehydrogenase: sequence, transcription anal
 A;Reference number: S18609; PMID:92114863; PMID:1766432
 A;Accession: S18609
 A;Molecule type: DNA
 A;Residues: 1-435 <BEN>
 A;Cross-references: UNIPROT:P29051; EMBL:X63837; NID:g49045; PIDN:CAM45327.1; PID:g49046
 A;Note: the source is designated as Halobacterium salinarum
 C;Superfamily: glutamate dehydrogenase (NAD(P)+)
 C;Keywords: NADP; oxidoreductase
 F;126/Binding site: substrate (Lys) #status predicted

Query Match 40.7%; Score 48; DB 1; Length 435;
 Best Local Similarity 41.2%; Pred. No. 13;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 GAPMNRWVCQMLETMF 19
 ||| ||| : ||| : ||| : ||| :
 Db 106 GLGMMWMTWKCAVMDLPF 122
 ||| ||| : ||| : ||| : ||| :

RESULT 5
 C83337
 hypothetical protein PA2465 [imported] - Pseudomonas aeruginosa (strain PA01)
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C;Accession: C83337
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brl
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A;Reference number: A82950; PMID:20437337; PMID:10984043
 A;Accession: C83337
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-382 <STO>
 A;Cross-references: UNIPROT:Q91117; GB:AE004674; GB:AE004091; NID:g9948512; PIDN:AAG0585
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: PA2465

Query Match 39.8%; Score 47; DB 2; Length 382;
 Best Local Similarity 60.0%; Pred. No. 16;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PRGAPMMWRW 10
 ||| ||| : ||| : ||| :
 Db 170 PRGRPFWSKW 179
 ||| ||| : ||| : ||| :

RESULT 6
 E72079
 phospholipase D homolog - Chlamydophila pneumoniae (strain CWL029)
 C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
 C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C;Accession: E72079
 R;Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999
 A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A;Reference number: A72000; PMID:99206606; PMID:10192388
 A;Accession: E72079
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-476 <ARN>
 A;Cross-references: UNIPROT:Q9Z8A8; GB:AE001626; GB:AE001363; NID:g4376708; PIDN:AAI857
 A;Experimental source: strain CWL029
 C;Genetics:
 A;Gene: CPH0435

Query Match 39.0%; Score 46; DB 2; Length 476;
 Best Local Similarity 41.7%; Pred. No. 28;
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 PMWRWVCQMLE 16
 ||| ||| : ||| : ||| :
 Db 364 PLWKKWFCEKIK 375
 ||| ||| : ||| : ||| :

RESULT 7
 AB6545
 phospholipase D type proteinase [imported] - Chlamydophila pneumoniae (strain J138)
 C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C;Accession: AB6545
 R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ishi
 Nucleic Acids Res. 28, 2311-2314, 2000
 A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A;Reference number: AB6491; PMID:20330349; PMID:10871362
 A;Accession: AB6545
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-476 <STO>

RESULT 11
AF2501
WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF2501
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N. 2003
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. 2001
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2501
A:Status: preliminary
A:Molecule type: DNA

A;Residues: 1-1191 <KUR>
A;Cross-references: UNIPROT:Q8YKV1; GB:BA000020; PIDN:BA078274.1; PID:g17135728; GSPDB:C
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr7190
A;Genome: plasmid

Query Match 38.1%; Score 45; DB 2; Length 1191;
Best Local Similarity 40.0%; Pred. No. 91;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 GAPMMRWVQMLET 17
Db 240 GQPELTQWLQCLINT 254
|||::|::|
|||::|::|

RESULT 12
AE3633
farnesyl cysteine carboxyl-methyltransferase BMEII0990 [imported] - Brucella melitensis
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AE3633
R;DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Majer, C.; Los, T.; Ivanova,
; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AE3633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-194 <KUR>
A;Cross-references: UNIPROT:Q8YBBO; UNIPROT:Q8FX32; GB:AE008918; PIDN:AAL54232.1; PID:g1
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII0990
A;Map position: II

Query Match 37.3%; Score 44; DB 2; Length 194;
Best Local Similarity 42.9%; Pred. No. 24;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 7 WMRVQCMLETMEL 20
Db 76 WNGWIGTLIELLFL 89
|||::|::|
|||::|::|

RESULT 13
AE2100
phosphonate ABC transport permease protein phnE [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AE2100
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; PMID:21595285; PMID:11759040
A;Accession: AE2100
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-264 <KUR>
A;Cross-references: UNIPROT:Q8YUJ0; GB:BA000019; PIDN:BA074055.1; PID:g17131448; GSPDB:C
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: phnE
C;Superfamily: phnE protein

Query Match 37.3%; Score 44; DB 2; Length 264;
Best Local Similarity 38.5%; Pred. No. 32;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 4 APMMRWVQCMLE 16
|||::|::|
|||::|::|

Db 102 APSWLQWLANILQ 114
|||::|::|
|||::|::|

RESULT 14
TI2260
caffeoyl-CoA O-methyltransferase (EC 2.1.1.104) - common ice plant (fragment)
C;Species: Mesembryanthemum crystallinum (common ice plant)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C;Accession: TI2260
R;Michalowski, C.B.; Bohnert, H.J.
submitted to the EMBL Data Library, March 1998
A;Description: Mesembryanthemum crystallinum caffeic acid 3-O-methyltransferase.
A;Reference number: Z17477
A;Accession: TI2260
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-350 <MTC>
A;Cross-references: UNIPROT:O65362; EMBL:AF067968; NID:g3176966; PID:g3176967
C;Function:
C;Superfamily: lignin biosynthesis
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase; S-adenosylmethionine

Query Match 37.3%; Score 44; DB 2; Length 350;
Best Local Similarity 41.7%; Pred. No. 42;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PRGAPMMRWVC 12
|||::|::|
|||::|::|

Db 242 PKGDAIFMKWIC 253
|||::|::|
|||::|::|

RESULT 15
S40146
catechol O-methyltransferase (EC 2.1.1.6) - cider tree
C;Species: Eucalyptus gunnii (cider tree)
C;Date: 19-May-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S40146
R;Grima-Pettenati, J.; Poeydomege, O.; Boudet, A.M.
submitted to the EMBL Data Library, August 1993
A;Reference number: S40146
A;Accession: S40146
A;Molecule type: mRNA
A;Residues: 1-366 <GRI>
A;Cross-references: UNIPROT:P46484; EMBL:X74814; NID:g437776; PID:g437777
C;Superfamily: O-methyltransferase
C;Keywords: methyltransferase; S-adenosylmethionine

Query Match 37.3%; Score 44; DB 2; Length 366;
Best Local Similarity 41.7%; Pred. No. 43;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PRGAPMMRWVC 12
|||::|::|
|||::|::|

Db 258 PKGDAIFMKWIC 269
|||::|::|
|||::|::|

Search completed: October 26, 2004, 06:46:24
Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:09:56 ; Search time 93 Seconds
(without alignments)

123.736 Million cell updates/sec

Title: US-10-066-965a-3

Perfect score: 118

Sequence: 1 PRGAPMWRVWVQCMLETHMFL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	44.9	185	2	Q61JH2
2	50.5	42.8	159	2	Q61699
3	50.5	42.8	307	2	Q82Q22
4	50	42.4	385	2	Q81DR5
5	50	42.4	459	2	Q73Y99
6	50	42.4	459	2	AAS04131
7	49	41.5	311	2	Q7U927
8	49	41.5	352	2	Q7RP82
9	49	41.5	429	1	DHE4_HALN1
10	49	41.5	545	1	AGAL_ASPNG
11	48	40.7	359	2	Q8G2S3
12	48	40.7	359	2	Q8YF17
13	48	40.7	435	1	DHE4_HALSA
14	48	40.7	612	2	Q8FTC3
15	48	40.7	616	2	Q8XEQ3
16	48	40.7	1424	1	DR21_ARATH
17	47.5	40.3	861	2	Q6E7I6
18	47	39.8	209	2	Q7UKY1
19	47	39.8	251	2	Q93HK1
20	47	39.8	260	2	Q82J76
21	47	39.8	382	2	Q91117
22	46	39.0	307	2	Q70IN1
23	46	39.0	307	2	CAB45100
24	46	39.0	310	2	Q8H727
25	46	39.0	476	2	Q928A8
26	46	39.0	499	2	Q7S1Z4
27	46	39.0	521	2	Q8MRN3
28	46	39.0	521	2	AAN09607
29	46	39.0	581	2	Q8T8M0
30	46	39.0	629	2	Q7W0A6
31	46	39.0	629	2	Q7W4N5

32 46 39.0 629 2 Q7WG61 bordetella
33 46 39.0 642 2 Q8IRU6 drosophila
34 46 39.0 662 2 Q8IRU4 drosophila
35 46 39.0 701 2 Q9W4S8 drosophila
36 46 39.0 701 2 AAF4S865 drosophila
37 46 39.0 777 1 CNA1_DROME P12252 drosophila
38 46 39.0 814 2 Q8IRU5 drosophila
39 46 39.0 814 2 AAN09605 AAN09605 drosophila
40 46 39.0 875 2 Q8FPG2 drosophila
41 46 39.0 903 2 Q9W4T0 drosophila
42 46 39.0 903 2 AAF4S862 drosophila
43 46 39.0 983 2 Q9W4S9 drosophila
44 46 39.0 983 2 AAF4S863 drosophila
45 46 39.0 1057 2 Q9W4T1 drosophila

ALIGNMENTS

RESULT 1

Q61JH2
ID Q61JH2 PRELIMINARY; PRT; 185 AA.
AC Q61JH2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HDC14914.
GN ORFNames=HDC14914;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14709175;
RA Hild M., Beckmann B., Haas S., Koch B., Solovyev V., Busold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hoheisel J., Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:R3-R3(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK002744; DAA04249.1; -
SQ SEQUENCE 185 AA; 21244 MW; 47401824148BA72D CRC64;

Query Match 44.9%; Score 53; DB 2; Length 185;

Best Local Similarity 75.0%; Pred. No. 4.7;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 WRRVWVQCMLETHM 18

Db 71 WRRVWVQCMLETHM 82

RESULT 2

Q61699
ID Q61699 PRELIMINARY; PRT; 159 AA.
AC Q61699;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative oxidoreductase (Fragment).
OS Streptomyces griseus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RA Kato J., Chi W., Ohnishi Y., Hong S., Horinouchi S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB182575; BAD24658.1; -
FT NON_TER 1 1

```
SQ SEQUENCE 159 AA; 17359 MW; 2D006A96A1A2B860 CRC64;
Query Match 42.8%; Score 50.5; DB 2; Length 159;
Best Local Similarity 56.2%; Pred. No. 9.8;
Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 4 APWMRWVCOMLETMF 19
Db 96 APPWLRW-CQPLRPLF 110

RESULT 3
Q82Q22 PRELIMINARY; PRT; 307 AA.
AC Q82Q22;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative oxidoreductase.
GN OrderedLocusNames=SAV699;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572949;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba I., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis; deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
DR EMBL; AP005023; BAC68409.1; -.
DR HSSP; O70351; IE3S.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; Adh_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; Beta_tubulin.
DR PRINTS; PR00081; GDHRD.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
DR PROSITE; PS00228; TUBULIN_B_AUTOREG; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 307 AA; 32843 MW; 2AE470395BAD60C4 CRC64;

Query Match 42.8%; Score 50.5; DB 2; Length 307;
Best Local Similarity 56.2%; Pred. No. 19;
Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 4 APWMRWVCOMLETMF 19
Db 249 APPWLRW-CQPLRPLF 263

RESULT 4
Q81DR5 PRELIMINARY; PRT; 385 AA.
ID Q81DR5
```

```
AC Q81DR5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Casein kinase II regulatory subunit, putative.
GN Name=PF13_0232;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52554.1; -.
DR HSSP; P13862; IQF8.
DR GO; GO:0005956; C:protein kinase CK2 complex; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008605; F:protein kinase CK2 regulator activity; IEA.
DR InterPro; IPR000704; CAS_Kinase_II.
DR Pfam; PF01214; CK_II_beta; 1.
DR PRINTS; PR00472; CAS_KINASEII.
DR ProDom; PD003829; CAS_Kinase_II; 1.
KW Kinase.
SQ SEQUENCE 385 AA; 45284 MW; 9972EB55B6EF87 CRC64;

Query Match 42.4%; Score 50; DB 2; Length 385;
Best Local Similarity 42.9%; Pred. No. 28;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 7 WMRWVCOMLETMFL 20
Db 161 WIEFQOLKQNLFL 174

RESULT 5
Q73YY9 PRELIMINARY; PRT; 459 AA.
AC Q73YY9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP1814;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017233; AAS04131.1; -.
DR InterPro; IPR009059; bi_resp_regltr_C.
DR InterPro; IPR009043; RNA_pol_sigma.
DR InterPro; IPR007627; Sigma70_r2.
DR InterPro; IPR007630; Sigma70_r4.
DR Pfam; PF04542; Sigma70_r2; 1.
DR Pfam; PF04545; Sigma70_r4; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 459 AA; 49141 MW; 5E3BD1C2DE326095 CRC64;

Query Match 42.4%; Score 50; DB 2; Length 459;
Best Local Similarity 63.8%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PRGAPMWRWV 11
Db 79 PTGRPAWRL 89
```

```

RESULT 6
AAS04131
ID AAS04131 PRELIMINARY; PRT; 459 AA.
AC AAS04131;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein.
GN MAP1814.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amons A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017233; AAS04131.1; -.
KW Hypothetical protein.
SQ SEQUENCE 459 AA; 49141 MW; 5B3BDLC2DE326095 CRC64;

Query Match 42.4%; Score 50; DB 2; Length 459;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PRGAPMWRWV 11
Db 79 PTGRPAWNRWL 89

RESULT 7
Q7U9Z7
ID Q7U9Z7 PRELIMINARY; PRT; 311 AA.
AC Q7U9Z7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=SYNW0105;
OS Synchococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synchococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569689; CAE06620.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000379; Ser esters.
DR Pfam; PF00561; Abhydrolase 1; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 311 AA; 34579 MW; E4D8440C13387B35 CRC64;

Query Match 41.5%; Score 49; DB 2; Length 311;
Best Local Similarity 57.1%; Pred. No. 32;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 PWRNRWVCQMLETM 18
Db 156 PAWMTWVRPLKTM 169

RESULT 8
Q7RP82
ID Q7RP82 PRELIMINARY; PRT; 352 AA.
AC Q7RP82;

```

```

DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Casein kinase II beta chain.
GN Name=PK01577;
OS Plasmodium yoelii yoelii.
OX Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteau M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shalom S.J., van Aken S.E., Riedmuller S.B., Feidblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000421; EAA20924.1; -.
DR GO; GO:0005956; C:protein kinase CK2 complex; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008605; F:protein kinase CK2 regulator activity; IEA.
DR InterPro; IPR000704; CAS_kinase_II.
DR Pfam; PF01214; CK_II_beta; 1.
DR PRINTS; PR00472; CASNKINASEII.
DR ProDom; PD003829; CAS_kinase_II; 1.
KW Kinase.
SQ SEQUENCE 352 AA; 41544 MW; 44A1799543D3B8BA CRC64;

Query Match 41.5%; Score 49; DB 2; Length 352;
Best Local Similarity 42.9%; Pred. No. 36;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 7 WMRWVCQMLETMFL 20
Db 126 WIENFCCQLKQNIPL 139

RESULT 9
DHE4_HALN1
ID DHE4_HALN1 STANDARD; PRT; 429 AA.
AC QHSM4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE NADP-specific glutamate dehydrogenase B (EC 1.4.1.4) (NADP-GDH B).
GN Name=gdhB; OrderedLocustNames=VNG0161G;
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OX Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Herquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";

```

Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- CATALYTIC ACTIVITY: L-glutamate + H(2)O + NADP(+) = 2-oxoglutarate
CC + NH(3) + NADPH.
CC -1- SUBUNIT: Homohexamer (By similarity).
CC -1- SIMILARITY: Belongs to the Glu/Leu/Phe/Val dehydrogenases family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE004982; AAG18779.1; -.
DR PIR; G84176; G84176.
DR HSP; P96110; I33B.
DR InterPro; IPR006095; GLFV dehydrog.
DR InterPro; IPR006096; GLFV dehydrog_C.
DR InterPro; IPR00205; NAD BS.
DR Pfam; PF02812; GLFV dehydrog; 1.
DR PRINTS; PR00082; GLFV dehydrog_N; 1.
DR PROSITE; PS00074; GLFV DEHYDROGENASE; 1.
KW Complete proteome; NADP; Oxidoreductase.
FT ACT_SITE 119 119 By similarity.
SQ SEQUENCE 429 AA; 45968 MW; A344F09C60B60A6 CRC64;

Query Match 41.5%; Score 49; DB 1; Length 429;
Best Local Similarity 41.2%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 GAPMWMRWVCOMLETMF 19
DB 99 GLSMWMTWKCAVMDLPP 115
ID _AGAL ASPNG STANDARD; PRT; 545 AA.
AC P28351;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Alpha-galactosidase A precursor (EC 3.2.1.22) (Melibiase).
GN Name:AGLA;
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-59.
RC STRAIN=ATCC 9089 / N402;
RX MEDLINE=92318893; PubMed=1320186;
RA den Herder I.F., Rosell A.M.M., van Zulen C.M., Punt P.J.,
RT "Cloning and expression of a member of the Aspergillus niger gene
RT family encoding alpha-galactosidase.";
RL Mol. Gen. Genet. 233:404-410(1992).
CC -1- FUNCTION: Represents a minor extracellular alpha-galactosidase
CC activity in A.niger.
CC -1- CATALYTIC ACTIVITY: Melibiose + H(2)O = galactose + glucose.
CC -1- PPM: A C-terminal Ser/Thr-rich region may provide possible sites
CC for O-glycosylation.
CC -1- SIMILARITY: Belongs to family 27 of glycosyl hydrolases.
CC -1- SIMILARITY: Contains 1 ricin B-type lectin domain.
CC -1- CAUTION: It is uncertain whether Met-1 or Met-9 is the initiator.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X63348; CAA44950.1; -.
DR PIR; S23582; S23582.
DR GlycoSuiteDB; P28351; -.
DR InterPro; IPR002241; Glyco hydro 27.
DR InterPro; IPR000111; Glyco hydro_GHD.
DR InterPro; IPR008997; RicinB_like_
DR InterPro; IPR000772; RicinB_lectin.
DR Pfam; PF02065; Melibiase; 1.
DR PRINTS; PR00652; RicinB_lectin; 3.
DR ProDom; PD002572; Glyco hydro_GHD; 1.
DR PROSITE; PS00512; ALPHA_GALACTOSIDASE; 1.
DR PROSITE; PS00231; RICIN_B_LLECTIN; 1.
KW Direct protein sequencing; Glycoprotein; Glycosidase; Hydrolase;
KW Lectin; Signal.
FT SIGNAL 1 31
FT CHAIN 32 545 Alpha-galactosidase A.
FT DOMAIN 421 518 Ricin B-type lectin.
FT ACT_SITE 162 162 Nucleophile (By similarity).
FT ACT_SITE 220 220 Proton donor (By similarity).
FT DISULFID 54 86 By similarity.
FT DISULFID 134 164 By similarity.
FT CARBOHYD 57 57 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 95 95 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 101 101 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 131 131 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 211 211 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 363 363 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 444 444 N-linked (GlcNAc. .) (Potential).
SQ SEQUENCE 545 AA; 60148 MW; 2DC5A710CE95D59C CRC64;

Query Match 41.5%; Score 49; DB 1; Length 545;
Best Local Similarity 42.1%; Pred. No. 55;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 PRGAPMWMRWVCOMLETMF 19
DB 43 PMGFNWARFMCIDNETLF 61
ID Q8G2S3 PRELIMINARY; PRT; 359 AA.
AC Q8G2S3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Oxidoreductase, Gfo/Idh/MocA family.
GN OrderedLocustNames=BR0241;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.L., Lindler L.E., Helling S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014337; AAN29190.1; -.
DR PIR; AH3465; AH3465.


```

DR TIGR; BR0241; ..
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000683; GFO/IDH/MocA_N.
DR Pfam; PF01408; GFO_IDH_MocA; 1.
KW Complete proteome.
SQ SEQUENCE 359 AA; 39386 MW; 21CA7CDB75AEB2E5 CRC64;

Query Match 40.7%; Score 48; DB 2; Length 359;
Best Local Similarity 45.8%; Pred. No. 52;
Matches 11; Conservative 2; Mismatches 3; Indels 8; Gaps 2;

Oy 1 PRGAP--MMMR-----WVCQMLE 16
||| | :| :| :| :|
Db 276 PRGPEDELWIRPKGGDDWVCVPLE 299
||| | :| :| :| :|

RESULT 12
O8YF17 PRELIMINARY; PRT; 359 AA.
AC O8YF17;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2003 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE OXIDOREDUCTASE [EC 1.1.1.-].
GN OrderedLocustNames=BMEI1710;
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OC NCBI_TaxID=29459;
RN EMBL; X63837; CAA45327.1; -.
RP STRAIN=16M / ATCC 23456 / Biotype 1;
RC MEDLINE=20020109; PubMed=1175688; DOI=10.1073/pnas.221575398;
RA Delvecchio V.G., Kapratral V., Redkar R.J., Patra G., Mujar C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kypides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RT Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009605; AAL52891.1; -.
DR PIR; AH3465; AH3465.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000683; GFO/IDH/MocA_N.
DR Pfam; PF01408; GFO_IDH_MocA; 1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 359 AA; 39386 MW; 21CA7CDB75AEB2E5 CRC64;

Query Match 40.7%; Score 48; DB 2; Length 359;
Best Local Similarity 45.8%; Pred. No. 52;
Matches 11; Conservative 2; Mismatches 3; Indels 8; Gaps 2;

Oy 1 PRGAP--MMMR-----WVCQMLE 16
||| | :| :| :| :|
Db 276 PRGPEDELWIRPKGGDDWVCVPLE 299
||| | :| :| :| :|

RESULT 13
DHE4_HALSA
ID DHE4_HALSA STANDARD; PRT; 435 AA.
AC P29051;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE NAD-specific glutamate dehydrogenase (EC 1.4.1.2) (NAD-GDH).
GN Name=gdhA;
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OC NCBI_TaxID=2242;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=CCM 2090;
RC MEDLINE=92114863; PubMed=1766432;
RA Benachenhou N., Baldacci G.;
RT "The gene for a halophilic glutamate dehydrogenase: sequence,
RT transcription analysis and phylogenetic implications.";
RL Mol. Gen. Genet. 230:345-352(1991).
[2]
RN SEQUENCE OF 119-128; 184-193 AND 268-276, AND CHARACTERIZATION.
RP PubMed=12052548;
RA Hayden B.M., Bonete M.J., Brown P.E., Moir A.J., Engel P.C.;
RT "Glutamate dehydrogenase of Halobacterium salinarum: evidence that the
RT gene sequence currently assigned to the NADP+-dependent enzyme is in
RT fact that of the NAD+-dependent glutamate dehydrogenase.";
RL FEMS Microbiol. Lett. 211:37-41(2002).
CC -!- CATALYTIC ACTIVITY: L-glutamate + H(2)O + NAD(+) = 2-oxoglutarate
CC + NH(3) + NADH.
CC -!- SUBUNIT: Homohexamer (By similarity).
CC -!- SIMILARITY: Belongs to the Glu/Leu/Phe/Val dehydrogenases family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X63837; CAA45327.1; -.
DR PIR; S18609; S18609.
DR HSSP; P96110; 1B26.
DR InterPro; IPR006095; GLFV dehydrog.
DR InterPro; IPR006096; GLFV dehydrog_C.
DR InterPro; IPR006097; GLFV dehydrog_N.
DR InterPro; IPR00205; NAD_BS.
DR Pfam; PF0208; GLFV dehydrog; 1.
DR Pfam; PF02812; GLFV dehydrog_N; 1.
DR PRINTS; PR00082; GLFV dehydrog_N; 1.
DR PROSITE; PS00074; GLFV dehydrog_N; 1.
DR PROSITE; PS00074; GLFV dehydrog_N; 1.
KW Direct protein sequencing; NAD; Oxidoreductase.
FT ACT SITE 126 126 By similarity.
SQ SEQUENCE 435 AA; 47458 MW; 0B4E940D2FF8B9D2 CRC64;

Query Match 40.7%; Score 48; DB 1; Length 435;
Best Local Similarity 41.2%; Pred. No. 63;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Oy 3 GAPMMRWVCQMLETFM 19
||| | :| :| :| :|
Db 106 GLGMMWTWKCAVMDLPF 122
||| | :| :| :| :|

RESULT 14
O8FTC3 PRELIMINARY; PRT; 612 AA.
ID O8FTC3
AC O8FTC3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ABC transporter TetB.
GN Name=tetB; OrderedLocustNames=CE1646;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OC NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314;
RX MEDLINE=22723752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:09:56 ; Search time 85.5 Seconds
(without alignments)

83.913 Million cell updates/sec

Title: US-10-066-965A-3

Perfect score: 118

Sequence: 1 PRGAPMMRWVCQMLETMFL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB ID	Description
1	49	41.5	441	2	AAR70209 A. niger
2	48.5	41.1	47	4	Abb38152 Peptide #
3	48.5	41.1	47	4	Aam31580 Peptide #
4	48.5	41.1	47	4	Abb23358 Protein #
5	48.5	41.1	47	4	Aam71301 Human Don
6	48.5	41.1	47	4	Aam58788 Human bra
7	48.5	41.1	47	4	Abg53012 Human liv
8	48.5	41.1	47	5	Abg41101 Human pep
9	48	40.7	96	3	Aay56647 Partial p
10	48	40.7	98	5	Aao21796 Lung-spec
11	48	40.7	140	3	Aay56712 Amino aci
12	47	39.8	318	4	Abg15007 Novel hum
13	47	39.8	421	7	Abc83694 Pseudomon
14	46.5	39.4	336	4	Abg08345 Novel hum
15	46	39.0	102	7	Adc72560 Human end
16	46	39.0	116	2	Aar66340 Human imm
17	46	39.0	196	2	Aay35036 Amino aci
18	46	39.0	476	5	Abb90563 Chlamydia
19	46	39.0	624	4	Abb59475 Drosophil
20	46	39.0	629	6	Abu23308 Protein e
21	45.5	38.6	718	8	Adm92105 S pneumon
22	45	38.1	129	5	Adk36943 Novel hum
23	45	38.1	136	4	Aau52383 Propionib
24	45	38.1	136	6	Abm48902 Propionib
25	45	38.1	1119	4	Abb63999 Drosophil

ALIGNMENTS

RESULT 1

AAR70209
ID AAR70209 standard; protein; 441 AA.

XX AC AAR70209;

XX 25-MAR-2003 (revised)

DT 21-SEP-1995 (first entry)

XX A. niger alpha-galactosidase.

DE Alpha-galactosidase; coffee bean; CBG; alpha-1,3-linked galactose;

KW B antigen; blood group.

XX Aspergillus niger.

XX WO9507088-A1.

XX 16-MAR-1995.

XX 26-AUG-1994; 94WO-US009662.

XX 08-SEP-1993; 93US-00118470.

XX (NYBL-) NEW YORK BLOOD CENT INC.

XX Zhu A, Goldstein J;

XX WPI; 1995-123231/16.

XX New recombinant coffee bean alpha-galactosidase - used for cleaving
PT alpha1,3-linked galactose residues on the surface of cells for prodn. of
PT blood prods.

XX Disclosure; Page 49-50; 67pp; English.

XX DNA encoding coffee bean alpha-galactosidase (CBG) was obtained by
PS isolating mRNA from coffee beans, prep. cDNA and subjecting this to PCR
XX amplification using primers based on the amino acid sequence of CBG. A
CC full-length cDNA clone is given in AAQ83523, which encoded a 42 kDa
CC protein (AAR70209). Recombinant CBG was produced in Sf9 cells. CBG showed
CC homology to alpha-galactosidases from guar, human placenta, S. cerevisiae
CC and A. niger (AAR70206-09). (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 441 AA;

Query Match 41.5%; Score 49; DB 2; Length 441;
Best Local Similarity 42.1%; Pred. No. 49;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 PRGAPMMRWVCOMLETMF 19
 Db 43 PMGFNWARFCDLNETLF 61

RESULT 2
 ABB38152
 ID ABB38152 standard; peptide; 47 AA.
 XX AC ABB38152;
 XX DT 04-FEB-2002 (first entry)
 XX DE Peptide #5658 encoded by human foetal liver single exon probe.
 XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX OS Homo sapiens.
 XX PN WO200157277-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US000669.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-08608408.
 XX PR 03-AUG-2000; 2000US-06632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human foetal liver.
 XX Claim 27; SEQ ID NO 30787; 639pp + Sequence listing; English.
 XX The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human foetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 47 AA;
 SQ Query Match 41.1%; Score 48.5; DB 4; Length 47;
 Best Local Similarity 60.0%; Pred. No. 6.6;
 Matches 12; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 PRGAPMMRWVCOMLETMF 20
 Db 13 PRGAPNLGAWVLCLE-IFL 31

RESULT 3
 AAM31580
 ID AAM31580 standard; protein; 47 AA.
 XX AC AAM31580;
 XX DT 17-OCT-2001 (first entry)

XX Peptide #5617 encoded by probe for measuring placental gene expression.
 DE Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 KW Homo sapiens.
 OS WO200157272-A2.
 PN 09-AUG-2001.
 XX PD 30-JAN-2001; 2001WO-US000663.
 XX PF 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-08608408.
 XX PR 03-AUG-2000; 2000US-06632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488897/53.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human placenta.
 XX Claim 27; SEQ ID NO 31849; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP;
 CC see AAI31315-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders
 XX Sequence 47 AA;
 SQ Query Match 41.1%; Score 48.5; DB 4; Length 47;
 Best Local Similarity 60.0%; Pred. No. 6.6;
 Matches 12; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 PRGAPMMRWVCOMLETMF 20
 Db 13 PRGAPNLGAWVLCLE-IFL 31

RESULT 4
 ABB23358
 ID ABB23358 standard; protein; 47 AA.
 XX AC ABB23358;
 XX DT 23-JAN-2002 (first entry)
 XX DE Protein #5357 encoded by probe for measuring heart cell gene expression.
 XX Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.
 XX Homo sapiens.
 OS WO200157274-A2.
 PN 09-AUG-2001.
 XX PD 30-JAN-2001; 2001WO-US000666.
 XX PF 30-JAN-2001; 2001WO-US000666.
 XX

PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX
XX WPI; 2001-488939/53.
XX DR
XX XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX PT
XX PS
XX PS Claim 15; SEQ ID NO 25128; 530pp; English.
XX XX
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart (see
XX CC ABA21535-ABAA11305). The present sequence is a protein encoded by one such
XX CC probe. The probes may be used for predicting, measuring and displaying
XX CC gene expression in samples derived from the human heart via microarrays.
XX CC By measuring gene expression, the probes are useful for predicting,
XX CC diagnosing, grading, staging, monitoring and prognosing diseases of the
XX CC human heart and vascular system e.g. cardiovascular disease,
XX CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX XX
XX XX Sequence 47 AA;
XX
Query Match 41.1%; Score 48.5; DB 4; Length 47;
Best Local Similarity 60.0%; Pred. No. 6.6;
Matches 12; Conservative 1; Mismatches 6; Indels 1; Gaps 1;
QY 1 PRGAPMMRWVCOMLETMFL 20
||||| ||| ||| :||
Db 13 PRGAPNLGAWVLQCLE-IFL 31
RESULT 5
AAM71301
ID AAM71301 standard; protein; 47 AA.
XX AC AAM71301;
XX XX
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31607.
XX XX
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD
XX XX
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX XX
XX XX WO200157276-A2.
XX PN
XX PD
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX
XX WPI; 2001-488900/53.
XX DR
XX XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human bone marrow.
XX PT
XX PS
XX PS Example 4; SEQ ID NO 31607; 650pp + Sequence Listing; English.
XX XX
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a
XX CC protein encoded by one of the probes of the invention
XX XX
XX XX Sequence 47 AA;
XX
Query Match 41.1%; Score 48.5; DB 4; Length 47;
Best Local Similarity 60.0%; Pred. No. 6.6;
Matches 12; Conservative 1; Mismatches 6; Indels 1; Gaps 1;
QY 1 PRGAPMMRWVCOMLETMFL 20
||||| ||| ||| :||
Db 13 PRGAPNLGAWVLQCLE-IFL 31
RESULT 6
AAM58788
ID AAM58788 standard; protein; 47 AA.
XX AC AAM58788;
XX XX
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30893.
XX XX
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD
XX XX
XX PF 09-AUG-2001.
XX PR
XX PR 30-JAN-2001; 2001WO-US000667.
XX PR
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX XX
XX XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX
XX WPI; 2001-483446/52.
XX DR
XX XX
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT brains.
XX XX
XX PS Example 4; SEQ ID NO 30893; 650pp + Sequence Listing; English.
XX XX
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX CC

CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention

XX Sequence 47 AA;

Query Match 41.1%; Score 48.5; DB 4; Length 47;
 Best Local Similarity 60.0%; Pred. No. 6.6;
 Matches 12; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

Qy 1 PRGAPMMRWVCOMLETMFL 20
 ||||| ||| :||
 Db 13 PRGAPNLGAWVLQCLE-IFL 31

RESULT 7

ABG53012
 ID ABG53012 standard; peptide; 47 AA.

XX AC ABG53012;

XX DT 25-FEB-2003 (first entry)

XX DE Human liver peptide, SEQ ID No 31660.

XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.

XX OS Homo sapiens.

XX PN WO200157273-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US0000664.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX PS WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human adult liver.

XX Claim 27; SEQ ID NO 31660; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 XX measuring human gene expression in a sample derived from human adult
 XX liver, comprising one of 13109 defined nucleotide sequences given in the
 XX specification (or complements/ fragments). The probe hybridises at high
 XX stringency to a nucleic acid molecule expressed in the human adult liver.
 XX (I) may be used for predicting, measuring and displaying gene expression
 XX in samples derived from human adult liver. The genes identified may be
 XX involved in genetic liver diseases such as cirrhosis,
 XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 XX associated with coronary heart disease. ABG47348-ABG59930 represent human
 XX liver single exon encoded peptides of the invention. Note: The sequence
 XX information for this patent does not appear in the printed specification
 XX but was obtained in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 47 AA;

Query Match

Best Local Similarity 41.1%; Score 48.5; DB 4; Length 47;

Matches 12; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

Matches

12; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

Qy 1 PRGAPMMRWVCOMLETMFL 20

||||| ||| :||

Db 13 PRGAPNLGAWVLQCLE-IFL 31

RESULT 8

ABG41101

ID ABG41101 standard; peptide; 47 AA.

XX AC ABG41101;

XX DT 19-AUG-2002 (first entry)

XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 30766.

XX KW Human; single exon probe; asthma; lung cancer; COPD; ILD;

XX KW chronic obstructive pulmonary disease; interstitial lung disease;

XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;

XX KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

XX KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX KW primary ciliary dyskinesia; pulmonary hypertension;

XX KW hyaline membrane disease.

XX OS Homo sapiens.

XX PN WO200186003-A2.

XX PD 15-NOV-2001.

XX PF 30-JAN-2001; 2001WO-US0000665.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX PS WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples.

XX Claim 27; SEQ ID NO 30766; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
 XX nucleic acid probes for measuring gene expression in a sample derived
 XX from human lung comprising single exon nucleic acid probes having one of
 XX 12614 nucleic acid sequences mentioned in the specification, or their
 XX complements or the 12387 open reading frames derived from the 12614
 XX probes. Also included are a microarray comprising the novel set of probes
 XX ; the novel set of probes which hybridise at high stringency to a nucleic
 XX acid expressed in the human lung; measuring gene expression in a sample
 XX derived from human lung, comprising (a) contacting the array with a
 XX collection of detectably labeled nucleic acids derived from human lung
 XX mRNA, and (b) measuring the label detectably bound to each probe of the
 XX array; identifying exons in a eukaryotic genome, comprising (a)
 XX algorithmically predicting at least one exon from genomic sequences of
 XX the eukaryote; and (b) detecting specific hybridisation of detectably
 XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 XX having a fragment identical to the predicted exon, the probe is included
 XX in the above mentioned microarray; assigning exons to a single gene,
 XX comprising (a) identifying exons from genomic sequence by the method
 XX above and (b) measuring the expression of each of the exons in several

CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a peptide/protein encoded by a single exon probe of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 XX
 SQ Sequence 47 AA;

Query Match 41.1%; Score 48.5; DB 5; Length 47;
 Best Local Similarity 60.0%; Pred. No. 6.6;
 Matches 12; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

Qy 1 PRGAPMWRWVCOMLETFEL 20
 Db 13 PRGAPNLGAWVLCLE-IFL 31

RESULT 9
 AAY56647
 ID AAY56647 standard; protein; 96 AA.

AC AAY56647;
 DT 15-FEB-2000 (first entry)
 DE Partial peptide fragment of chimpanzee VH cDNA clone 41-9.

XX Complementarity determining region; antibody; primate; immunogenicity;
 KW Old World ape; Old World monkey; antigen-binding affinity.
 XX Pan troglodytes.

XX WO9955369-A1.
 PN 04-NOV-1999.

XX 28-APR-1999; 99WO-US009131.
 XX 28-APR-1999; 99US-0083367P.

XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX Taylor AH;

XX WPI; 2000-023265/02.
 DR N-PSDB; AAZ39315.

XX Antibodies containing donor complementarity determining regions and non-
 PT human primate acceptor frameworks, having reduced immunogenicity in
 PT humans.

PS Claim 20; Page 55; 123pp; English.

XX The invention provides an antibody (Ab) comprising donor CDRs
 CC (complementarity determining regions) derived from a non-human antigen-
 CC specific donor antibody, and an acceptor framework from a non-human
 CC primate. The Abs are prepared by grafting CDRs from a non-human antigen-
 CC specific donor antibody onto homologous Old World ape or monkey acceptor
 CC frameworks. The Abs have reduced immunogenicity and are better tolerated

CC in humans (because of the close similarity between the human and primate
 CC proteins), but retain the full antigen-binding affinity of the donor
 CC antibody

SQ Sequence 96 AA;

Query Match 40.7%; Score 48; DB 3; Length 96;
 Best Local Similarity 87.5%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 WMRWVCOM 14
 Db 33 WMGWVCOM 40

RESULT 10
 AAO21796
 ID AAO21796 standard; protein; 98 AA.

XX AAO21796;

XX 13-SEP-2002 (first entry)

XX Lung-specific amino acid sequence SEQ ID No 107.

XX Cytostatic; lung specific protein; metastasis; lung cancer; gene therapy;
 KW transgenic animal; analysis algorithm; lung.

XX Homo sapiens.

XX WO200246224-A2.

XX 13-JUN-2002.

XX 26-OCT-2001; 2001WO-US050687.

XX 26-OCT-2000; 2000US-0243461P.

XX 20-NOV-2000; 2000US-0252055P.

XX 22-NOV-2000; 2000US-0252496P.

XX (DIAD-) DIADEXUS INC.

XX Recipon H, Sun Y, Chen S, Liu C, Turner LR;

XX WPI; 2002-537558/57.

XX Lung specific polypeptides useful for detecting, diagnosing, monitoring,
 PT treating, staging and predicting cancers in humans with cancer and non-
 PT cancerous lung disease.

PS Claim 11; Page 210; 228pp; English.

XX The invention relates to an isolated lung specific polypeptide comprising
 CC an amino acid sequence 60% identical to a sequence chosen from 61
 CC sequences such as 68, 63, 569, 51, 105, 711, 49 or 77 amino acids, or an
 CC amino acid sequence encoded by a nucleic acid molecule comprising a
 CC sequence chosen from 81 sequences such as 463, 554, 1074, 1277, 676, 524,
 CC 258 or 4106 bp given in the specification. The isolated polypeptide and
 CC its encoding polynucleotide are useful for diagnosing and monitoring the
 CC presence and metastases of lung cancer in a patient, by determining an
 CC amount of the polypeptide or polynucleotide in a sample of a patient, and
 CC comparing it to the amount of lung specific marker in a normal control,
 CC where a difference in the amount of the nucleic acid or the polypeptide
 CC in the sample compared to that of normal control is associated with
 CC presence of lung cancer. The polypeptide and polynucleotide are also
 CC useful for producing engineered lung tissue for treatment and research.
 CC The polypeptide and polynucleotide are useful as components in databases
 CC useful for search analyses and in sequence analysis algorithms. The
 CC polynucleotide is useful for producing transgenic animals and cells and
 CC also in gene therapy. This sequence represents a lung-specific amino acid
 CC sequence relating to the invention

XX Sequence 98 AA;

PF 23-DEC-2002; 2002WO-US041612.
XX
PR 21-DEC-2001; 2001US-0342756P.
XX
PA (DIAD-) DIADEXUS INC.
XX
FI Sun Y, Liu C;
XX
DR WPI; 2003-577666/54.
XX
PT Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX
PS Claim 12; SEQ ID NO 600; 824pp; English.
XX
CC The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present amino acid sequence represents a
CC human endometrial specific protein of the invention.
XX
SQ Sequence 102 AA;

Query Match 39.0%; Score 46; DB 7; Length 102;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

QY 2 RCAPMMWR---WVCO 13
|||:|
DB 65 RGAPVWPERDTLWLCQ 80

Search completed: October 26, 2004, 06:45:13
Job time : 89.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:45:25 ; Search time 64.25 Seconds
(without alignments)
100.781 Million cell updates/sec

Title: US-10-066-965A-3

Perfect score: 118

Sequence: 1 PRGAPMMRWVCOMLETMFL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PTCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	118	100.0	20	14	US-10-066-965A-3
2	118	100.0	20	14	US-10-066-965A-22
3	96	81.4	20	14	US-10-066-965A-4
4	96	81.4	20	14	US-10-066-965A-23
5	56	47.5	173	15	US-10-425-114-60423
6	50.5	42.8	123	9	US-09-771-161A-171
7	50.5	42.8	307	14	US-10-156-761-8240
8	50	42.4	87	16	US-10-437-963-19829
9	49	41.5	429	14	US-10-369-493-18537
10	49	41.5	1451	16	US-10-437-963-175423
11	48.5	41.1	47	9	US-09-864-761-38656
12	48	40.7	96	9	US-09-905-243-15
13	48	40.7	98	14	US-10-034-934-107

14	47	39.8	63	15	US-10-424-599-204224	Sequence 204224, A
15	47	39.8	260	14	US-10-156-761-10442	Sequence 10442, A
16	47	39.8	358	16	US-10-437-963-133725	Sequence 133725, A
17	46	39.0	196	15	US-10-289-762-454	Sequence 454, App
18	46	39.0	476	15	US-10-312-273-75	Sequence 75, Appl
19	46	39.0	584	9	US-09-802-741A-4	Sequence 4, Appl
20	46	39.0	629	15	US-10-282-122A-51232	Sequence 51232, A
21	45	38.1	37	16	US-10-437-963-196778	Sequence 196778, A
22	45	38.1	211	15	US-10-424-599-169363	Sequence 169363, A
23	45	38.1	267	16	US-10-437-963-167292	Sequence 167292, A
24	45	38.1	356	15	US-10-425-114-63867	Sequence 63867, A
25	45	38.1	583	16	US-10-437-963-188631	Sequence 188631, A
26	45	38.1	913	16	US-10-437-963-188636	Sequence 188636, A
27	44.5	37.7	116	16	US-10-437-963-127057	Sequence 127057, A
28	44.5	37.7	155	16	US-10-437-963-162418	Sequence 162418, A
29	44	37.3	74	15	US-10-424-599-184426	Sequence 184426, A
30	44	37.3	308	15	US-10-108-260A-3067	Sequence 3067, Ap
31	44	37.3	33	15	US-10-424-599-161201	Sequence 161201, A
32	44	37.3	368	9	US-09-796-256A-6	Sequence 6, Appl
33	44	37.3	376	9	US-09-882-872-2	Sequence 2, Appl
34	44	37.3	376	14	US-10-184-648-84	Sequence 84, Appl
35	44	37.3	376	15	US-10-296-606-2	Sequence 2, Appl
36	44	37.3	413	13	US-10-106-534-6	Sequence 6, Appl
37	44	37.3	414	13	US-10-006-867-102	Sequence 102, App
38	44	37.3	414	13	US-10-052-586-314	Sequence 314, App
39	44	37.3	414	13	US-10-063-547-102	Sequence 102, App
40	44	37.3	414	13	US-10-063-551-102	Sequence 102, App
41	44	37.3	414	14	US-10-174-590-314	Sequence 314, App
42	44	37.3	414	14	US-10-176-758-314	Sequence 314, App
43	44	37.3	414	14	US-10-175-737-314	Sequence 314, App
44	44	37.3	414	14	US-10-063-616-102	Sequence 102, App
45	44	37.3	414	14	US-10-174-581-314	Sequence 314, App

ALIGNMENTS

RESULT 1
US-10-066-965A-3
; Sequence 3, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-3

Query Match	100.0%	Score 118;	DB 14;	Length 20;
Best Local Similarity	100.0%	Pred. No. 4.8e-10;		
Matches	20;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;
QY	1	PRGAPMMRWVCOMLETMFL 20		
Db	1	PRGAPMMRWVCOMLETMFL 20		
RESULT 2				
US-10-066-965A-22				
; Sequence 22, Application US/10066965A				
; Publication No. US20030143626A1				

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-23

Query Match      81.4%; Score 96; DB 14; Length 20;
Best Local Similarity 85.0%; Pred. No. 7e-07;
Matches 17; Conservative 1; Mismatches 2; Indels 0;

Qy 1 PRGAPMWRVVCQMLETFL 20
   |||||:|||||:
Db 1 PRGAPMWRVVCQMLETFL 20
   |||||:|||||:

RESULT 5
US-10-425-114-60423
; Sequence 60423, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Ass
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60423
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3592-081-A11_FLI.pep
US-10-425-114-60423

Query Match      47.5%; Score 56; DB 15; Length 173;
Best Local Similarity 66.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 4; Indels 0;

Qy 1 PRGAPMWRVVC 12
   |||||:|
Db 53 PTGAPMWRVWC 64
   |||||:|

RESULT 6
US-09-771-161A-171
; Sequence 171, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 171
; LENGTH: 123

```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-171

Query Match      42.8%; Score 50.5; DB 9; Length 123;
Best Local Similarity 42.9%; Pred. No. 12;
Matches          9; Conservative 1; Mismatches 2; Indels 9; Gaps 1;

QY   1 PRG-----APMMRWVC 12
    |||
Db   79 PRGHKISDYFETAPLWFRWC 99
    |||||

RESULT 7
US-10-156-761-8240
; Sequence 8240, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8240
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8240

Query Match      42.8%; Score 50.5; DB 14; Length 307;
Best Local Similarity 56.2%; Pred. No. 28;
Matches          9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY   4 APMMRWVCOMLETMF 19
    |||.:.|||.:.|
Db   249 APPWLRLW-CQPLRPLF 263
    |||.:.|||.:.|

RESULT 8
US-10-437-963-199829
; Sequence 199829, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 199829
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_7326C.1.pep
US-10-437-963-199829

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-171

Query Match      42.4%; Score 50; DB 16; Length 87;
Best Local Similarity 41.2%; Pred. No. 11;
Matches          7; Conservative 6; Mismatches 2; Indels 2; Gaps 1;

QY   3 GAPMMRWVCOMLETMF 19
    |||.:.|||.:.|
Db   69 GAPVW--WTCELVEVVW 83
    |||.:.|||.:.|

RESULT 9
US-10-369-493-18537
; Sequence 18537, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18537
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Halobacterium sp. NRC-1
US-10-369-493-18537

Query Match      41.5%; Score 49; DB 14; Length 429;
Best Local Similarity 41.2%; Pred. No. 61;
Matches          7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY   3 GAPMMRWVCOMLETMF 19
    |||.:.|||.:.|
Db   99 GLSMWMTWKCAVMDLPF 115
    |||.:.|||.:.|

RESULT 10
US-10-437-963-175423
; Sequence 175423, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 175423
; LENGTH: 1451
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_7326C.1.pep
US-10-437-963-175423

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-171

Query Match      42.8%; Score 50.5; DB 9; Length 123;
Best Local Similarity 42.9%; Pred. No. 12;
Matches          9; Conservative 1; Mismatches 2; Indels 9; Gaps 1;

QY   1 PRG-----APMMRWVC 12
    |||
Db   79 PRGHKISDYFETAPLWFRWC 99
    |||||

RESULT 7
US-10-156-761-8240
; Sequence 8240, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8240
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8240

Query Match      42.8%; Score 50.5; DB 14; Length 307;
Best Local Similarity 56.2%; Pred. No. 28;
Matches          9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY   4 APMMRWVCOMLETMF 19
    |||.:.|||.:.|
Db   249 APPWLRLW-CQPLRPLF 263
    |||.:.|||.:.|

RESULT 8
US-10-437-963-199829
; Sequence 199829, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 199829
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_7326C.1.pep
US-10-437-963-199829
```

```
Best Local Similarity 41.2%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 GAPMMRWVCOMLETMF 19
    ||| ||| ||| :
Db 908 GLPLWCLWALQMLRSWY 924

RESULT 11
US-09-864-761-38656
; Sequence 38656, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeotica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38656
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006024.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
```

```
; OTHER INFORMATION: EST_HUMAN HIT: EP241410.1, EVALUE 6.00e-01
US-09-864-761-38656

Query Match 41.1%; Score 48.5; DB 9; Length 47;
Best Local Similarity 60.0%; Pred. No. 10;
Matches 12; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 PRGAPMMRWVCOMLETMF 20
    ||| ||| ||| |||
Db 13 PRGAPNUGAVLQCLE-IFL 31

RESULT 12
US-09-905-243-15
; Sequence 15, Application US/09905243
; Patent No. US20020062009A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Alexander H
; TITLE OF INVENTION: Monoclonal Antibodies with Reduced
; FILE REFERENCE: P50770
; CURRENT APPLICATION NUMBER: US/09/905,243
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/300,970
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (31)...(35)
; OTHER INFORMATION: CDRI
; NAME/KEY: DOMAIN
; LOCATION: (50)...(66)
; OTHER INFORMATION: CDRII
US-09-905-243-15

Query Match 40.7%; Score 48; DB 9; Length 96;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 WMRWVCOM 14
    ||| ||| |||
Db 33 WMGWVCOM 40

RESULT 13
US-10-034-934-107
; Sequence 107, Application US/10034934
; Publication No. US20030068624A1
; GENERAL INFORMATION:
; APPLICANT: Recipon, Herve
; APPLICANT: Sun, Yongming
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific
; FILE REFERENCE: DEX-0245
; CURRENT APPLICATION NUMBER: US/10/034,934
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 60/243,461
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/252,055
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/252,496
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 107
```

; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-034-934-107

Query Match 40.7%; Score 48; DB 14; Length 98;
Best Local Similarity 41.2%; Pred. No. 23;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 PRGAPMMRWVCQMLET 17
|||:|:|:|:
Db 80 PGGAHLWFFWLCDRVT 96

RESULT 14

US-10-424-599-204224
; Sequence 204224, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 204224
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_26441C.1.pap
US-10-424-599-204224

Query Match 39.8%; Score 47; DB 15; Length 63;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 RGAPMMRWVCQ 13
||:|:|:|:
Db 35 RPVPLWSRWCCR 46

RESULT 15

US-10-156-761-10442
; Sequence 10442, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIEA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10442
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10442

Query Match 39.8%; Score 47; DB 14; Length 260;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 PRGAPMMRWVC 12
|||:|:|:
Db 68 PRGAPRWPRAGVC 79

Search completed: October 26, 2004, 07:10:19
Job time : 65.25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:20:00 ; Search time 22 Seconds
(without alignments)
60.289 Million cell updates/sec

Title: US-10-066-965A-3

Perfect score: 118

Sequence: 1 PRGAPMMRWVCOMLETWFL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	42.4	447	3	US-09-070-356-7
2	48	40.7	315	4	US-09-248-796A-20438
3	47	39.8	421	4	US-09-252-991A-32440
4	46	39.0	116	3	US-08-545-809A-134
5	46	39.0	136	4	US-09-198-452A-454
6	46	39.0	584	2	US-08-987-466-4
7	46	39.0	584	3	US-09-240-359-4
8	45	38.1	115	4	US-09-270-767-45700
9	44	37.3	368	3	US-08-991-677-6
10	43	36.4	1138	4	US-09-252-991A-24340
11	43	36.4	1336	4	US-09-328-352-5500
12	42	35.6	91	4	US-09-270-767-38667
13	42	35.6	91	4	US-09-270-767-53884
14	42	35.6	364	3	US-08-860-820-2
15	42	35.6	430	4	US-08-311-731A-170
16	41	34.7	26	3	US-08-525-539A-24
17	41	34.7	237	4	US-09-270-767-37845
18	41	34.7	237	4	US-09-270-767-53062
19	41	34.7	257	1	US-08-287-959-4
20	41	34.7	283	4	US-09-270-767-38878
21	41	34.7	283	4	US-09-270-767-54095
22	41	34.7	600	3	US-08-904-871-2
23	41	34.7	600	3	US-08-904-871-3
24	41	34.7	766	5	PCT-US94-00198-6
25	41	34.7	1129	3	US-08-904-871-11
26	41	34.7	1323	1	US-08-026-138E-4
27	40.5	34.3	539	4	US-09-614-912-144

28	40	33.9	62	4	US-09-270-767-61766	Sequence 61766, A
29	40	33.9	109	4	US-09-270-767-37621	Sequence 37621, A
30	40	33.9	109	4	US-09-270-767-52838	Sequence 52838, A
31	40	33.9	137	4	US-09-248-796A-18260	Sequence 18260, A
32	40	33.9	139	4	US-09-355-925-7	Sequence 7, Appli
33	40	33.9	139	4	US-09-355-925-8	Sequence 8, Appli
34	40	33.9	139	4	US-09-269-921-105	Sequence 105, App
35	40	33.9	139	4	US-09-269-921-108	Sequence 108, App
36	40	33.9	139	4	US-09-269-921-109	Sequence 109, App
37	40	33.9	139	4	US-09-269-921-110	Sequence 110, App
38	40	33.9	139	4	US-09-269-921-111	Sequence 111, App
39	40	33.9	139	4	US-09-269-921-112	Sequence 112, App
40	40	33.9	139	4	US-09-269-921-113	Sequence 113, App
41	40	33.9	139	4	US-09-269-921-114	Sequence 114, App
42	40	33.9	139	4	US-09-269-921-115	Sequence 115, App
43	40	33.9	139	4	US-09-269-921-116	Sequence 116, App
44	40	33.9	139	4	US-09-269-921-117	Sequence 117, App
45	40	33.9	139	4	US-09-269-921-118	Sequence 118, App

ALIGNMENTS

RESULT 1
US-09-070-356-7
; Sequence 7, Application US/09070356
; Patent No. 6228631
; GENERAL INFORMATION:
; APPLICANT: Alex Zhu
; APPLICANT: Jack Goldstein
; TITLE OF INVENTION: Recombinant a-N-
; TITLE OF INVENTION: Acetylalactosaminidase
; TITLE OF INVENTION: Enzyme and cDNA Encoding
; TITLE OF INVENTION: Said Enzyme
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amster, Rothstein & Ebenstein
; STREET: 90 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,356
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: 08/037,248
; APPLICATION NUMBER: March 26, 1993
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; REFERENCE/DOCKET NUMBER: 63475/12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447
; TYPE: amino acid
; STRANDEDNESS: Double
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to mRNA
; HYPOTHETICAL: no
; ANTI-SENSE: yes
; FRAGMENT TYPE:

ORIGINAL SOURCE:
ORGANISM: Aspergillus niger
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE: library
POSITION IN GENOME: unknown
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: Aspergillus niger a-galactosidase
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: den Herder et al
TITLE: Cloning and Expression of a Member
of the Aspergillus niger Gene Family
TITLE: Encoding a-Galactosidase
JOURNAL: Molecular and General Genetics
VOLUME: 233
PAGES: 404-410
DATE: 1992
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:

US-09-070-356-7

Query Match 42.4%; Score 50; DB 3; Length 447;
Best Local Similarity 42.1%; Pred. No. 9;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 PRGAPMWRWVCOMLETFM 19
DB 43 PMGPNWAWFMCDLNETLF 61

RESULT 2

US-09-248-796A-20438
Sequence 20438, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 20438
LENGTH: 315
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-20438

Query Match 40.7%; Score 48; DB 4; Length 315;
Best Local Similarity 35.0%; Pred. No. 13;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 PRGAPMWRWVCOMLETFM 20
DB 104 PHNPMWLKRWYFIISMVF 123

RESULT 3

US-09-252-991A-32440
Sequence 32440, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32440
LENGTH: 421
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32440

Query Match 39.8%; Score 47; DB 4; Length 421;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PRGAPMWRW 10
DB 209 PRGRPFWSKW 218

RESULT 4

US-08-545-809A-134
Sequence 134, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid

QY 1 PRGAPMWRWVCOMLETFM 19
DB 43 PMGPNWAWFMCDLNETLF 61

US-09-248-796A-20438
Sequence 20438, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 20438
LENGTH: 315
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-20438

Query Match 40.7%; Score 48; DB 4; Length 315;
Best Local Similarity 35.0%; Pred. No. 13;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 PRGAPMWRWVCOMLETFM 20
DB 104 PHNPMWLKRWYFIISMVF 123

Tue Oct 26 08:02:22 2004

us-10-066-965a-3.ra1

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-545-809A-134
Query Match 39.0%; Score 46; DB 3; Length 116;
Best Local Similarity 70.0%; Pred. No. 9.1;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 WMRWVCQMLE 16
   |||||
Db 51 WMRWVCQAPE 60

RESULT 5
US-09-198-452A-454
; Sequence 454, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 454
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-454

Query Match 39.0%; Score 46; DB 4; Length 196;
Best Local Similarity 41.7%; Pred. No. 16;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 PNMWRWVCQMLE 16
   |||||
Db 84 PLWKWFCFKL 95

RESULT 6
US-08-987-466-4
; Sequence 4, Application US/08987466
; Patent No. 5922595
; GENERAL INFORMATION:
; APPLICANT: Fisher, Douglas A.
; APPLICANT: Gooding, Doug
; APPLICANT: Streeter, Dave
; TITLE OF INVENTION: CYCLIC-GMP PHOSPHODIESTERASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/987,466
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0442 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 829179
US-08-987-466-4

Query Match 39.0%; Score 46; DB 3; Length 584;
Best Local Similarity 43.8%; Pred. No. 48;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 5 PNMWRWVCQMLETMTPL 20
   |||||
Db 414 PLYKRWVALLMEFFL 429

RESULT 7
US-09-240-359-4
; Sequence 4, Application US/09240359
; Patent No. 6255456
; GENERAL INFORMATION:
; APPLICANT: Fisher, Douglas A.
; APPLICANT: Gooding, Doug
; APPLICANT: Streeter, Dave
; TITLE OF INVENTION: CYCLIC-GMP PHOSPHODIESTERASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/240,359
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/987,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0442 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 829179
US-09-240-359-4

Query Match 39.0%; Score 46; DB 3; Length 584;
Best Local Similarity 43.8%; Pred. No. 48;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 5 PWRWVCOMLETMFL 20
 ::|||::|||
 Db 414 PLYKRVALIMEEFL 429

RESULT 8
 US-09-270-767-45700
 ; Sequence 45700, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 45700
 ; LENGTH: 115
 ; TYPE: PRT
 ; ORGANISM: *Drosophila melanogaster*
 US-09-270-767-45700

Query Match 38.1%; Score 45; DB 4; Length 115;
 Best Local Similarity 46.2%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 PRGAPMMRWVCQ 13
 |||||
 Db 6 PRPQPPWAEWACR 18

RESULT 9
 US-08-991-677-6
 ; Sequence 6, Application US/08991677A
 ; Patent No. 6252135
 ; GENERAL INFORMATION:
 ; APPLICANT: Chiang, Vincent L
 ; APPLICANT: Carraway, Daniel T
 ; APPLICANT: Smeltzer, Richard H
 ; TITLE OF INVENTION: Production of Syringyl Lignin in *Gymnosperms*
 ; FILE REFERENCE: 50617
 ; CURRENT APPLICATION NUMBER: US/08/991,677A
 ; CURRENT FILING DATE: 1997-12-16
 ; EARLIER APPLICATION NUMBER: US 60/033,381
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 368
 ; TYPE: PRT
 ; ORGANISM: *Liquidambar styraciflua*
 US-08-991-677-6

Query Match 37.3%; Score 44; DB 3; Length 368;
 Best Local Similarity 41.7%; Pred. No. 60;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 PRGAPMMRWVC 12
 |||||
 Db 259 PKGDAIFMKWIC 270

RESULT 10
 US-09-252-991A-24340
 ; Sequence 24340, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 24340
 ; LENGTH: 1118
 ; TYPE: PRT
 ; ORGANISM: *Pseudomonas aeruginosa*
 US-09-252-991A-24340

Query Match 36.4%; Score 43; DB 4; Length 1118;
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PRGAPMMWR 9
 |||||
 Db 489 PRGSPAWLR 497

RESULT 11
 US-09-328-352-5500
 ; Sequence 5500, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; CURRENT APPLICATION NUMBER: GTC99-03PA
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 5500
 ; LENGTH: 1336
 ; TYPE: PRT
 ; ORGANISM: *Acinetobacter baumannii*
 US-09-328-352-5500

Query Match 36.4%; Score 43; DB 4; Length 1336;
 Best Local Similarity 33.3%; Pred. No. 3.2e+02;
 Matches 7; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 1 PRGAPMMWR--WVCOMLETMFL 19
 |||||
 Db 1186 PRSAPMTMFWLSEGTSLY 1206

RESULT 12
 US-09-270-767-38667
 ; Sequence 38667, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 38667
 ; LENGTH: 91
 ; TYPE: PRT
 ; ORGANISM: *Drosophila melanogaster*
 US-09-270-767-38667

Query Match 35.6%; Score 42; DB 4; Length 91;
 Best Local Similarity 75.0%; Pred. No. 29;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAPMMW 8

```
; MOLECULE TYPE: protein
US-08-860-820-2
Query Match 35.6%; Score 42; DB 3; Length 364;
Best Local Similarity 46.7%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PRGAPMMRWVCQML 15
Db 243 PRGKDNMRWIASMV 257

RESULT 15
US-08-311-731A-170
; Sequence 170, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 170:
; LENGTH: 430 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium leprae
; US-08-311-731A-170

Query Match 35.6%; Score 42; DB 4; Length 430;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 RGAPMMRW 10
Db 10 RGMPPQWR 18

Search completed: October 26, 2004, 06:47:57
Job time : 23 secs
```

```
; MOLECULE TYPE: protein
US-09-270-767-53884
Query Match 35.6%; Score 42; DB 4; Length 91;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRGAPMMW 8
Db 65 PRGAPVWV 72

RESULT 14
US-08-860-820-2
; Sequence 2, Application US/08860820
; Patent No. 6245967
; GENERAL INFORMATION:
; APPLICANT: Sonnewald, Uwe
; APPLICANT: Kossmann, Jens
; APPLICANT: Bowien, Botho
; TITLE OF INVENTION: PROCESS AND DNA MOLECULES FOR INCREASING
; TITLE OF INVENTION: THE PHOTOSYNTHESIS RATE IN PLANTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/860,820
; FILING DATE: 04-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19502053.7
; FILING DATE: 13-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: AGREVO-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:19:49 ; Search time 17 Seconds
(without alignments)
113.196 Million cell updates/sec

Title: US-10-066-965A-4

Perfect score: 115

Sequence: 1 PRGAPMLRCVQMLETXFL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: Pirl:*

2: Pirl2:*

3: Pirl3:*

4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	43.5	1044	T43800	protein kinase bub
2	47	40.9	773	JU0135	aldehyde dehydroge
3	44	38.3	413	T29505	hypothetical prote
4	44	38.3	1483	S42839	tl6G12.5 protein -
5	44	38.3	26926	I38344	titin, cardiac mus
6	42.5	37.0	622	AG0001	potassium transpor
7	42	36.5	92	B82066	phosphocarrier pro
8	42	36.5	299	S68198	probable transcrip
9	42	36.5	322	H87131	probable transcrip
10	42	36.5	542	S58102	hypothetical prote
11	42	36.5	545	S23592	alpha-galactosidas
12	42	36.5	570	T59402	NGFI-A-binding pro
13	42	36.5	826	T43638	caspase-related pr
14	42	36.5	2410	T43731	cell wall alpha-gl
15	41	35.7	122	G69532	conserved hypotet
16	41	35.7	165	H95957	probable C4-dicarb
17	41	35.7	190	S46926	phytochrome - Mars
18	41	35.7	231	S58931	H+-transporting tw
19	41	35.7	434	H70989	probable GABA perm
20	41	35.7	463	T02481	probable protein k
21	41	35.7	466	A36904	prochlorophyllid
22	41	35.7	484	JC7350	N-acetylglucosamin
23	41	35.7	486	JC7351	N-acetylglucosamin
24	41	35.7	499	E86206	hypothetical prote
25	41	35.7	618	T28668	hypothetical prote
26	41	35.7	618	G59432	MacGAP protein [im
27	41	35.7	769	D86678	hypothetical prote
28	41	35.7	786	T08664	Toll protein-like
29	41	35.7	992	A39931	protein-tyrosine k

protein-tyrosine k
Flt3 protein - mou
phytochrome / prot
genome polyprotein
nonstructural poly
replicase polyprot
hypothetical prote
hypothetical prote
conserved hypotet
hypothetical prote
membrane associate
pyruvate dehydroge
transketolase NMB1
transketolase (EC
homeobox protein (

30 41 35.7 993 2 A36873
31 41 35.7 1000 2 S18827
32 41 35.7 1303 1 S27396
33 41 35.7 2164 1 GNNY89
34 41 35.7 2479 1 MNWY8A
35 40.5 35.2 1759 2 T03725
36 40 34.8 253 2 T51913
37 40 34.8 318 2 T24116
38 40 34.8 389 2 A11506
39 40 34.8 390 2 A11147
40 40 34.8 432 2 T37509
41 40 34.8 452 2 B97095
42 40 34.8 617 2 H75540
43 40 34.8 659 2 B81082
44 40 34.8 659 2 A81862
45 40 34.8 745 2 D96829

ALIGNMENTS

RESULT 1

T43800
protein kinase bub1, checkpoint-associated - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43800; T40943
R:Bernard, P.; Hardwick, K.; Javerzat, J.P.
J. Cell Biol. 143, 1775-87, 1998
A:Title: Fission yeast bub1 is a mitotic centromere protein essential for the spindle che
A:Reference number: Z14026; MUID:9864354; PMID:9864354
A:Accession: T43800
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1044 <BB>
A:Cross-references: UNIPROT:O94751; EMBL:AF064796; PIDN:AAC98348.1
A:Experimental source: strain sp011
R:Lucas, M.; Gaillardin, C.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 1999
A:Reference number: Z21959
A:Accession: T40943
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1044 <LUC>
A:Cross-references: EMBL:AL035259; PIDN:CAA22865.1; GSPDB:GN00068; SPDB:SPCC1322.12c
A:Experimental source: strain 972h-; cosmid c1322
C:Genetics:
A:Gene: bub1; SPCC1322.12c
A:Map position: 3
C:Function:
A:Description: essential for the fission yeast spindle checkpoint response to spindle dan

Query Match 43.5%; Score 50; DB 2; Length 1044;

Best Local Similarity 53.3%; Pred. No. 7.2;

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 6 MLCRCVQMLETXFL 20

Db 52 VWRCIEWLLETXFL 66

RESULT 2

JU0135
aldehyde dehydrogenase (pyrroloquinoline-quinone) (EC 1.2.99.3) precursor - Acetobacter s
N:Alternate names: aldehyde dehydrogenase (acceptor)
C:Species: Acetobacter sp.
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 12-Jul-2004
C:Accession: JU0135
R:Tamaki, T.; Horinouchi, S.; Fukaya, M.; Okumura, H.; Kawamura, Y.; Beppu, T.
J. Biochem. 106, 541-544, 1989
A:Title: Nucleotide sequence of the membrane-bound aldehyde dehydrogenase gene from Acet
A:Reference number: JU0135; MUID:90110042; PMID:2606906
A:Accession: JU0135

A:Molecule type: DNA
A:Residues: 1-773 <TAM>
A:Cross-references: UNIPROT:PI7201
C:Comment: This enzyme is a membrane-bound protein.
C:Superfamily: Alcohol and aldehyde dehydrogenases are two enzymes in Acetobacter species which
C:KeyWords: isooquinoline 1-oxidoreductase beta subunit
F:1-44/Domain: signal sequence #status predicted <SIG>
F:45-773/Product: aldehyde dehydrogenase #status experimental <MAT>

Query Match 40.9%; Score 47; DB 2; Length 773;
Best Local Similarity 61.5%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 RGAPWLRVCVQM 14
DB 631 RGMPTWDRCAQV 643

RESULT 3
T29505
hypothetical protein F43H9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29505
R:Wu, X.; Le, T.
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid F43H9.
A:Reference number: Z20629
A:Accession: T29505
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-413 <WUX>
A:Cross-references: UNIPROT:Q20377; EMBL:U50307; PIDN:AAA92305.1; CESP:F43H9.3
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F43H9.3
A:Introns: 32/1; 138/2; 206/3; 292/1; 379/1

Query Match 38.3%; Score 44; DB 2; Length 413;
Best Local Similarity 46.2%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 MWLRVCVQMLETK 18
DB 216 LWLRACDKLEVR 228

RESULT 4
S42839
T16G12.5 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S42839
R:Thomas, K.
submitted to the EMBL Data Library, February 1994
A:Reference number: S42837
A:Accession: S42839
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1483 <THO>
A:Cross-references: UNIPROT:Q22528; EMBL:Z30317; NID:G457454; PID:G457457
C:Genetics:
A:Introns: 91/3; 127/3; 161/1; 282/3; 345/2; 474/3; 515/2; 538/1; 555/3; 782/1; 979/1; 1
C:Superfamily: Caenorhabditis elegans T16G12.5 protein

Query Match 38.3%; Score 44; DB 2; Length 1483;
Best Local Similarity 50.0%; Pred. No. 88;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 FMWLRVCVQMLETKFL 20
DB 750 FMWLKVCAGKQMTKIL 765

RESULT 5

I38344
titin, cardiac muscle [validated] - human
N:Alternate names: connectin
C:Contains: serine/threonine-specific protein kinase (EC 2.7.1.1.-)
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C:Accession: I38344; I38345; S20898; S20899; S63665; S37393
R:Labeit, S.; Kolmerer, B.
Science 270, 293-296, 1995
A:Title: Titins: Giant proteins in charge of muscle ultrastructure and elasticity.
A:Reference number: A57430; MUID:96026330; PMID:7569978
A:Accession: I38344
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A:Molecule type: mRNA
A:Residues: 1-26926 <LAB1>
A:Cross-references: UNIPROT:Q10466; EMBL:X90568; NID:G1017424; PID:G1017425
R:Musco, G.; Tziatzios, C.; Schuck, P.; Pastore, A.
Biochemistry 34, 553-561, 1995
A:Title: Dissecting titin into its structural motifs: identification of an alpha-helix
A:Reference number: I38345; MUID:95119041; PMID:7819249
A:Accession: I38345
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1977-2014 <MUS>
A:Cross-references: EMBL:X83270; NID:G602579; PIDN:CAA58243.1; PID:G602580
R:Labeit, S.; Gautel, M.; Lakey, A.; Trinick, J.
EMBO J. 11, 1711-1716, 1992
A:Title: Towards a molecular understanding of titin.
A:Reference number: S20897; MUID:92258380; PMID:1582406
A:Accession: S20898
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 13597-14200, 'I', 14202-14696 <LAB2>
A:Cross-references: EMBL:X64698; NID:G37192; PIDN:CAA45939.1; PID:G37193
A:Accession: S20897
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 16330-16382, 'S', 16384-16756, 'F', 16758-16860 <LAB3>
A:Cross-references: EMBL:X64699; NID:G37190; PIDN:CAA45940.1; PID:G37191
A:Accession: S20899
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 'P', 22278-22431, 'R', 22433-22448, 'G', 22450-22453, 'Q', 22455-22480, 'TR', 22483-2
A:Cross-references: EMBL:X64697; NID:G37190; PIDN:CAA45938.1; PID:G37195
R:Kolmerer, B.; Olivieri, N.; Witt, C.C.; Hertrmann, B.G.; Labeit, S.
J. Mol. Biol. 256, 556-563, 1996
A:Title: Genomic organization of M line titin and its tissue-specific expression in two
A:Reference number: S63665; MUID:96177761; PMID:8604138
A:Accession: S63665
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 26729-26825 <KOL>
A:Cross-references: EMBL:X92412; NID:G1236761
R:Gautel, M.; Leonard, K.; Labeit, S.
EMBO J. 12, 3827-3834, 1993
A:Title: Phosphorylation of KSP motifs in the C-terminal region of titin in differentia
A:Reference number: S37393; MUID:94008990; PMID:8404852
A:Accession: S37393
A:Molecule type: mRNA
A:Residues: 26831-26926 <GAU>
R:Improt, S.; Politou, A.S.; Pastore, A.
submitted to the Brookhaven Protein Data Bank, February 1996
A:Reference number: A66736; PDB:1TIT
A:Contents: annotation; conformation by (1)H-NMR, residues 5253-5341
R:Pfuhl, M.; Pastore, A.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A66201; PDB:1NCT
A:Contents: annotation; conformation by (1)H-NMR, residues 'S', 26059-26155
C:Genetics:

A:Gene: GDB:TTN
A:Cross-references: GDB:127867; OMIM:188840
A:Map position: 2q31-2q32
C:Function:
A:Description: structural protein forming filaments in striated muscle
C:Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; pro
C:Keywords: alternative splicing; calmodulin binding; cardiac muscle; duplication; glyco
structural protein
F:24752-25008/Domain: protein kinase homology <XIN>
F:84,177,905,2276,2378,2459,2481,2563,2669,2763,2896,3088,3179,3384,3432,3628,3772,4068
98,11066,11488,11515,11635,11949,12170,12478,12526,12645,12875,13001,13036,13295,13540,1
F:16780,16976,17579,17602,17667,17681,17845,17899,18121,18188,18209,18336,18670,18680,18
21900,21935,22295,22495,22627,22897,23024,23318,23883,24012,24177,24290,24447,24642,248
F:26171,26178,26184,26190/Binding site: phosphate (Ser) (covalent) #status experimental
Query Match 38.3%; Score 44; DB 1; Length 26926;
Best Local Similarity 44.4%; Pred. No. 1.2e+03;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
Qy 1 PRGAPMWLRCVCOMLETK 18
Db 7131 PRGSKWVACGEPVAETK 7148
RESULT 6
AG0001
potassium transport protein kup [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG0001
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0001
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-622 <KUR>
A:Cross-references: UNIPROT:Q8ZJTO; GB:AL590842; PIDN:CAC88873.1; PID:gl5978121; GSPDB:G
C:Genetics:
A:Gene: kup
Query Match 37.0%; Score 42.5; DB 2; Length 622;
Best Local Similarity 41.7%; Pred. No. 70;
Matches 10; Conservative 3; Mismatches 6; Indels 5; Gaps 1;
Qy 1 PRGAPMWLRC-----VCOMLETK 19
Db 546 PNVAEIFHRCGLGSLCOMMETSF 569
RESULT 7
B82066
phosphocarrier protein Npr VC2533 [imported] - Vibrio cholerae (strain N16961 serogroup
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: B82066
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: B82066
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-92 <HEI>
A:Cross-references: UNIPROT:Q9KP46; GB:AE004322; GB:AE003852; NID:g9657104; PIDN:AAF9567
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:

A:Gene: VC2533
A:Map position: 1
Query Match 36.5%; Score 42; DB 2; Length 92;
Best Local Similarity 58.3%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Qy 8 LRCVCOMLETKF 19
Db 77 LQAVCQLIEQKF 88
RESULT 8
S68198
probable transcription regulator (carotenoid gene cluster orf 11) - Myxococcus xanthus
C:Species: Myxococcus xanthus
C:Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S68198; S67958
R:Botella, J.; Murrillo, F.; Ruiz-vazquez, R.
submitted to the EMBL Data Library, November 1994
A:Description: A cluster of structural and regulatory genes for light-induced carotenogen
A:Reference number: S68198
A:Accession: S68198
A:Molecule type: DNA
A:Residues: 1-299 <BOT>
A:Cross-references: UNIPROT:Q50900; EMBL:Z21955; NID:g577589; PIDN:CRA79965.1; PID:g5775
R:Botella, J.A.; Murrillo, F.J.; Ruiz-Vazquez, R.
Eur. J. Biochem. 233, 238-248, 1995
A:Title: A cluster of structural and regulatory genes for light-induced carotenogenesis i
A:Reference number: S67950; MUID:96061955; PMID:7588751
A:Accession: S67958
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 7-68 <BOW>
A:Cross-references: EMBL:Z21955
C:Keywords: DNA binding
Query Match 36.5%; Score 42; DB 2; Length 299;
Best Local Similarity 34.8%; Pred. No. 44;
Matches 8; Conservative 4; Mismatches 3; Indels 8; Gaps 1;
Qy 1 PRGAPMWL-----RCVCOML 15
Db 260 PRGLPVWVGGRRRSHQAVCERL 282
RESULT 9
H87131
probable transcription regulator [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: H87131
R:Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hor
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: H87131
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-322 <STO>
A:Cross-references: UNIPROT:Q9CBN3; GB:AL450380; NID:gl3093502; PIDN:CAC30736.1; GSPDB:G
C:Genetics:
A:Gene: ML1783
Query Match 36.5%; Score 42; DB 2; Length 322;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 2 RGAPMWLRC 10
:|||||:|

A:Accession: T43435
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1848, 'A', 1850-2027, 'C', 2029-2410 <KAT>
 A:Cross-references: EMBL:AB019183; NID:G3808057; PIDN:BAA34054.1; PID:G3808058
 R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.
 submitted to the EMBL Data Library, June 1998
 A:Reference number: Z22013
 A:Accession: T41740
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-255, 'Y', 257-1210 <LYN>
 A:Cross-references: EMBL:AL023781; NID:G6272239; PIDN:CAA19332.1; PID:G3184076; GSPDB:GN
 A:Experimental source: strain 972h-; cosmid c338
 R:Volckaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, January 1999
 A:Reference number: Z21957
 A:Accession: T40920
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1392-2410 <VOL>
 A:Cross-references: EMBL:AL035218; NID:G4160576; PIDN:CAA22822.1; PID:G4160577; GSPDB:GN
 A:Experimental source: strain 972h-; cosmid c1281
 R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
 DNA Res. 4, 363-369, 1997
 A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
 A:Reference number: Z17323; MUID:98162722; PMID:9501991
 A:Accession: T43003
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 2044-2410 <VOS>
 A:Cross-references: EMBL:D89202; NID:G1749611; PIDN:BAA13863.1; PID:G1749612
 A:Experimental source: strain PR45
 R:Hochstenbach, F.; Klausner, R.D.
 submitted to the EMBL Data Library, May 1998
 A:Reference number: Z22659
 A:Accession: T43746
 A:Status: translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 2076-2410 <HOC>
 A:Cross-references: EMBL:AF063305; NID:G3387871; PIDN:AAC39519.1; PID:G3387872
 A:Experimental source: strain 972
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Saunders, D.; Harris, D.
 submitted to the EMBL Data Library, March 1999
 A:Reference number: Z21972
 A:Accession: T41143
 A:Status: translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1157-1465 <WOO>
 A:Cross-references: EMBL:AL049472; NID:G4538675; PIDN:CAB39330.1; PID:G4538676; GSPDB:GN
 A:Experimental source: strain 972h-; cosmid c17A7
 C:Genetics:
 A:Gene: ags1; mok1; SPDB:SPCC338.01c; SPDB:SPCC1281.01; SPDB:SPCC17A7.01
 A:Map position: 3
 C:Function:
 A:Description: involved in cell morphogenesis independently of the actin cytoskeleton; d
 C:Keywords: Glycosyltransferase; hexosyltransferase; transmembrane protein

 Query Match 36.5%; Score 42; DB 2; Length 2410;
 Best Local Similarity 45.5%; Pred. No. 2.8e+02;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

 Qy 4 APMWLRVCQM 14
 ||| : |||
 Db 2382 APFWNIICQL 2392

 RESULT 15
 G69532
 conserved hypothetical protein AF2263 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: G69532

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.A
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: G69532
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-122 <KLE>
 A:Cross-references: UNIPROT:O28021; GB:AE000948; GB:AE000782; NID:G2689271; PIDN:AAB8898;
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF2263

Query Match 35.7%; Score 41; DB 2; Length 122;
 Best Local Similarity 44.4%; Pred. No. 28;
 Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 2 RGAPMWLRVCQMLETKF 19
 :|||:| :|||:
 Db 38 KGAPVSLSAIQKMLRREP 55

Search completed: October 26, 2004, 06:46:25
 Job time : 18 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:09:56 ; Search time 93 Seconds

(without alignments)
123.736 Million cell updates/sec

Title: US-10-066-965A-4

Perfect score: 115

Sequence: 1 PRGAPWLRVCQMLETKFL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	43.5	1044	BUB1 SCHPO	O94751 schizosacch
2	48	41.7	579	Q7TSE3	Q7tse3 mus musculus
3	47	40.9	188	Q9FSA3	Q9fsa3 stangeria e
4	47	40.9	773	DHAQ ACEPO	PI7201 acetobacter
5	47	40.9	913	Q90XB9	Q90xb9 gallus gall
6	46	40.0	307	Q70IN1	Q70in1 pseudomonas
7	46	40.0	307	CAE45100	Cae45100 pseudomon
8	45	40.0	773	Q30329	Q30329 acetobacter
9	45	39.1	120	Q8FPA6	Q8fpa6 corynebacte
10	45	39.1	251	Q93HK1	Q93hk1 streptomyc
11	45	39.1	260	Q82J76	Q82j76 streptomyc
12	45	39.1	371	Q8BP03	Q8bp03 mus musculu
13	45	39.1	441	Q7RY30	Q7ry30 nemirospora
14	45	39.1	486	Q7R032	Q7r032 enericella
15	45	39.1	663	Q8K005	Q8k005 mus musculu
16	44	38.3	140	Q863H0	Q863h0 macaca mula
17	44	38.3	243	Q93FT4	Q93ft4 cowdria rum
18	44	38.3	256	Q9F1T4	Q9f1t4 arabidopsis
19	44	38.3	348	Q91TR9	Q91tr9 tupaia herp
20	44	38.3	413	Q20377	Q20377 caenorhabdi
21	44	38.3	485	NAB1 MESAU	Q35589 mesorichetu
22	44	38.3	1142	Q22528	Q22528 caenorhabdi
23	44	38.3	26926	Q10466	Q10466 homo sapien
24	44	38.3	26926	Q8WZB3	Q8wzb3 homo sapien
25	44	38.3	34350	Q8WZ42	Q8wz42 homo sapien
26	43.5	37.8	193	Q8Z9V3	Q8z9v3 oryza sativ
27	43.5	37.8	193	Q8Z9V3	Bac98592 oryza sat
28	43.5	37.8	193	BAC959593	Bac95953 oryza sat
29	43.5	37.8	272	Q6NK48	Q6nk48 corynebacte
30	43.5	37.8	272	CAE48694	Cae48694 corynebac
31	43.5	37.8	376	Q7MWT7	Q7mwt7 porphyromon

32	37.4	76	2	Q9L374	Q9l374 rhizobium l
33	37.4	160	2	Q9EB09	Q9eb09 sesbania mo
34	37.4	291	2	Q7GNQ2	Q7gnq2 giardia lam
35	37.4	297	2	Q92L10	Q92l10 rhizobium m
36	37.4	311	2	Q8TW68	Q8tw68 methanopyru
37	37.4	409	2	Q854Q1	Q854q1 mycobacteri
38	37.4	451	2	Q8VUY2	Q8vuy2 drosophila
39	37.4	482	2	Q8CJZ5	Q8cjz5 streptomyce
40	37.4	546	2	Q6VZNI	Q6vzn1 oryza sativ
41	37.4	546	2	BAD03478	Bad03478 oryza sat
42	37.4	546	2	BAD03752	Bad03752 oryza sat
43	37.4	645	2	Q82R39	Q82r39 streptomyce
44	42.5	219	2	Q9W5K9	Q9w5k9 euphorbia e
45	42.5	479	2	Q84SV9	Q84sv9 oryza sativ

ALIGNMENTS

RESULT 1

BUB1 SCHPO

ID BUB1 SCHPO STANDARD; PRT; 1044 AA.

AC Q94751;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Checkpoint serine/threonine-protein kinase bub1 (EC 2.7.1.37).
GN Name=bub1; ORFNames=SPCC1322.12c;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]

SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.

RC STRAIN=SP011;

RX MEDLINE=99085059; PubMed=9864354;

RT "Fission yeast Bubi is a mitotic centromere protein essential for the

spindle checkpoint and the preservation of correct ploidy through

mitosis."

RL J. Cell Biol. 143:1775-1787(1998).

[2]

SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

SGourcos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,

Rutherford S., Rutter S., Saunders D., Seeger K., Sharp S.,

Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,

Taylor K., Tavey R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,

Welters J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,

Borzum K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,

Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,

Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe.";

RL Nature 415:871-880(2002).

CC -!- FUNCTION: Involved in cell cycle checkpoint enforcement. Acts to

RA Tamaki T., Horinouchi S., Fukaya M., Okumura H., Kawamura Y.,
RA Beppu T.;
RT "Nucleotide sequence of the membrane-bound aldehyde dehydrogenase gene
from Acetobacter polyoxogenes.";
RL J. Biochem. 106:541-544(1989).
CC -!- CATALYTIC ACTIVITY: An aldehyde + acceptor + H(2)O = a carboxylate
CC + reduced acceptor.
CC -!- COFACTOR: PQQ.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; D00521; BAA00408.1; -;
CC PIR; JU0135; JU0135.
CC InterPro; IPR008274; Aldxan_dh_bind.
CC InterPro; IPR006311; Tat.
CC Pfam; PF02738; Ald_xan_dh_C2; 2.
CC TIGRFAMs; TIGR01409; TAT_signal_seg; 1.
CC Direct protein sequencing; Inner membrane; Oxidoreductase; PQQ;
CC Signal; Transmembrane.
CC SIGNAL 1 44
CC CHAIN 45 773 Membrane-bound aldehyde dehydrogenase
CC [pyrroloquinoline-quinone].
CC SEQUENCE 773 AA; 84249 MW; 93A76757B3B054FC CRC64;
CC -----
Query Match 40.9%; Score 47; DB 1; Length 773;
Best Local Similarity 61.5%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 RGAPMWLRVCVCOM 14
DB 631 RGMPTWDRCAQV 643

RESULT 5
Q90XB9 PRELIMINARY; PRT; 913 AA.
AC Q90XB9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Patched 2 (Fragment).
GN Name=PTC2;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Parse R.V. Jr., Vogan K.J., Tabin C.J.;
RT "Ptc1 and Ptc2 Transcripts Provide Distinct Readouts of Hedgehog
RT Signaling Activity during Chick Embryogenesis.";
RL Dev. Biol. 0:0-0(2001).
DR EMBL; AF409095; AAK97655.1; -;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR004766; Patchedtm_recept.
DR InterPro; IPR000731; SSD_5TM.
CC Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00918; 2A060602; 1.
DR PROSITE; PS50156; SSD; 1.
FT NON_TER 913
SQ SEQUENCE 913 AA; 101047 MW; DBD8117B5D842F33 CRC64;
Query Match 40.9%; Score 47; DB 2; Length 913;
us-10-066-965a-4.rup

Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 RGAPMWLRVCVCOML 15
DB 26 RGAPLWLRARFOAL 39

RESULT 6
Q70IN1 PRELIMINARY; PRT; 307 AA.
AC Q70IN1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative transcriptional repressor.
GN Name=paax2;
OS Pseudomonas sp. Y2.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=198931;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y2;
RA Bartolome-Martin D., Martinez-Garcia E., Mascaraque V., Rubio J.,
RA Perera J., Alonso S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AJ579894; CAE45100.1; -;
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
SQ SEQUENCE 307 AA; 34884 MW; 72348C9C6E3CA676 CRC64;
Query Match 40.0%; Score 46; DB 2; Length 307;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 PRGAPMWLRVCVCOMLE 16
DB 36 PRGGTVWLGSLIQLE 51

RESULT 7
CAE45100 PRELIMINARY; PRT; 307 AA.
AC CAE45100;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative transcriptional repressor.
GN PAAX2.
OS Pseudomonas sp. Y2.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=198931;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y2;
RA Bartolome-Martin D., Martinez-Garcia E., Mascaraque V., Rubio J.,
RA Perera J., Alonso S.;
RT "Characterization of a second functional gene cluster for the
RT catabolism of phenylacetic acid in Pseudomonas sp. strain Y2.";
RL Submitted (AUG-2003) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AJ579894; CAE45100.1; -;
DR Hypothetical protein; Porin.
SQ SEQUENCE 307 AA; 34884 MW; 72348C9C6E3CA676 CRC64;
Query Match 40.0%; Score 46; DB 2; Length 307;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 PRGAPMWLRVCVCOMLE 16
DB 36 PRGGTVWLGSLIQLE 51

```

KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 120 AA; 13296 MW; 9C8D11681057EAD0 CRC64;

Query Match 39.1%; Score 45; DB 2; Length 120;
Best Local Similarity 40.0%; Pred. No. 39;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 PRGAPMWLRVCVQMLETKEFL 20
DB ||||| : : : : :
49 PRDAQFWRDCLARLLHRKVL 68

RESULT 10
Q93HK1 PRELIMINARY; PRT; 251 AA.
AC Q93HK1;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Phosphopantetheinyl transferase.
GN Name=pptA;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; AB070940; BAB69186.1; -.
DR GO; GO:0003366; C:enterobactin synthetase complex; IEA.
DR GO; GO:0008897; F:holo-[acyl-carrier protein] synthase activity; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0016740; E:transferase activity; IEA.
DR GO; GO:0003239; E:enterobactin biosynthesis; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR008278; 4-PPT synthetD.
DR Pfam; PF01648; ACPS; 1.
DR PRINTS; PR01399; ENT5NTHTASED.
DR Transferase.
KW SEQUENCE 251 AA; 26956 MW; 6B774EDC2B573E98 CRC64;

Query Match 39.1%; Score 45; DB 2; Length 251;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 PRGAPMWLRVCV 12
DB ||||| : : : : :
59 PRGAPRWPAQVC 70

RESULT 11
Q82J76 PRELIMINARY; PRT; 260 AA.
AC Q82J76;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative phosphopantetheinyl transferase.
GN Name=pptA; OrderedLocusNames=SAV2905;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RESULT 8
O30329 PRELIMINARY; PRT; 773 AA.
AC O30329;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Aldehyde dehydrogenase precursor.
GN Name=aldH;
OS Acetobacter europaeus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconacetobacter.
OX NCBI_TaxID=33995;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DES11, and DSM 6160;
RX MEDLINE=97385093; PubMed=9238099;
RA Turner C.A., Vela C., Thony-Meyer L., Meile L., Teuber M.;
RT "Biochemical and genetic characterization of the acetaldehyde de
RT hydrogenase complex from Acetobacter europaeus.";
RL Arch. Microbiol. 168:81-91(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DES11, and DSM 6160;
RA Turner C.A.K.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y08696; CAA69955.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR008274; Aldkan_dh_bind.
DR InterPro; IPR006311; Tat.
DR Pfam; PF02738; Ald_Xan_dh_C2; 2.
DR TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
KW SIGNAL.
FT SIGNAL 1 44 Potential.
FT CHAIN 45 773 Potential.
SQ SEQUENCE 773 AA; 84105 MW; 3B549B655BE63F56 CRC64;

Query Match 40.0%; Score 46; DB 2; Length 773;
Best Local Similarity 53.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 RGAPMWLRVCQM 14
DB ||||| : : : : :
631 RGMPTWIGCAQV 643

RESULT 9
Q8FP46 PRELIMINARY; PRT; 120 AA.
AC Q8FP46;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=CEI1942;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314;
RX MEDLINE=22723752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
DR EMBL; AP005220; BAC18752.1; -.

```



```

RC STRAIN=WA-4680;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., T.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis; deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=WA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL: AP005013; BAC70616.1; -.
DR GO: GO:0009366; C:enterobactin synthetase complex; IEA.
DR GO: GO:0008977; F:holo-[acyl-carrier protein] synthase activity; IEA.
DR GO: GO:0000287; F:magnesium ion binding; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0009239; P:enterobactin biosynthesis; IEA.
DR GO: GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro: IPR008278; 4-PPT_transf.
DR InterPro: IPR003542; Embac_synthetD.
DR Pfam: PF01648; ACPS; 1.
DR PRINTS: PR01399; ENTSTNTHASES.
DR Complete proteome; Transferase.
DR KW SEQUENCE 260 AA; 28003 MW; ED3E123F1B969A4 CRC64;
SQ
Query Match 39.1%; Score 45; DB 2; Length 260;
Best Local Similarity 66.7%; Pred. No. 81;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 PRGAPMWLRVCV 12
Db 68 PRGAPRWPAQVC 79
|||||
|||

RESULT 12
Q8BF03 ID QBEP03 PRELIMINARY; PRT; 371 AA.
AC QBEP03;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 12 days embryo female mullerian duct includes surrounding
DE region cDNA, RIKEN full-length enriched library, clone:6820430N10
DE product:similar to MACGAP (Fragment).
DE Name:Arhgap18;
GN Name=Arhgap18;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;

```

```

RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;
RX MEDLINE=2030913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nishikawa K., Katsunai T., Tashiro H., Itoh M.,
RA Konno H., Akiyama J., Nishi K., Hazama M., Nishine T., Harada A.,
RA Sumi N., Ishii Y., Nakamura S., Ikegami T., Kashiwagi K.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Izawa M., Ohara E., Watahiki M.,
RA Fujiwaka S., Inoue K., Iogawa Y., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;
RX STRAIN=C57BL/6J; TISSUE=Mullerian duct includes surrounding region;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaiura S., Takeda Y., Tanaka Y.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
EMBL: AK078522; BAC37322.1; -.
MGD: MGI:1921160; Arhgap18.
DR InterPro: IPR000198; RhoGAP.
DR InterPro: IPR008936; Rho GAP.
DR Pfam: PF00620; RhoGAP; 1.
DR SMART: SM00324; RhoGAP; 1.
DR PROSITE: PS50238; RHO GAP; 1.
FT NON TER 1
FT SEQUENCE 371 AA; 42337 MW; D8B81555C1DDB70B CRC64;
SQ
Query Match 39.1%; Score 45; DB 2; Length 371;
Best Local Similarity 52.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 GAPMWLRVCVCMLETKF 19
Db 76 GAAMRIKMLCQLEAKF 92
|||||
|||

RESULT 13
Q7RY30 ID Q7RY30 PRELIMINARY; PRT; 441 AA.
AC Q7RY30;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU04534.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

```

131 PRSAPLWLR 139

DB

RESULT 15

Q8KQ05

ID Q8K005 PRELIMINARY; PRT; 663 AA.

AC Q8K005;

DT 01-OCT-2002 (T-EMBLrel. 22, Created)

DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)

DE Rho GTPase activating protein 18.

GN Name=Arhgap18;

GN Name=Arhgap18;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

XX [1]

RY SEQUENCE FROM N.A.

RP STRAIN=FVB/N; TISSUE=Liver;

RP MEDLINE=22388257; PubMed=12477932;

RX STRAUSBERG R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Schmutz J., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP SEQUENCE FROM N.A.

RP STRAIN=FVB/N; TISSUE=Liver;

RA Strausberg R.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

RL EMBL; BC030858; AAH30858.1; -

DR HSSP; Q98935; 1F7C.

DR MGD; MGI:1321160; Arhgap18.

DR InterPro; IPR000196; RhoGAP.

DR InterPro; IPR008936; Rho GAP.

DR Pfam; PF00620; RhoGAP; 1.

DR SMART; SM00324; RhoGAP; 1.

DR PROSITE; PS50238; RHO GAP; 1.

SR SEQUENCE 663 AA; 74930 MW; BB9ED06B850FEF41 CRC64;

SQ

Query Match 39.1%; Score 45; DB 2; Length 663;

Best Local Similarity 52.9%; Pred. No. 1.9e+02;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 GAPMMLRCVCOMLETKF 19

DB 368 GAAMRIKLCQLEAKF 384

Job time : 93 secs

Search completed: October 26, 2004, 06:39:23

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:09:56 ; Search time 85.5 Seconds
(without alignments)
83.913 Million cell updates/sec

Title: US-10-066-965A-4

Perfect score: 115

Sequence: 1 PRGAPWLRVCVQMLETKFL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 segs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	43.5	112	4	Aao09931 Human pol
2	49	42.6	78	4	Aao09705 Human pol
3	47	40.9	773	1	Aap91653 Membrane-
4	46	40.0	88	4	Aau31146 Novel hum
5	46	40.0	274	7	Adc87493 Human GPC
6	46	40.0	634	4	Abg01589 Novel hum
7	45.5	39.6	195	7	Adc87493 Human GPC
8	45.5	39.6	264	7	Adc87493 Human GPC
9	45.5	39.6	482	4	Abg25840 Novel hum
10	45	39.1	32	4	Aam21627 Peptide #
11	45	39.1	32	4	Abb43988 Peptide #
12	45	39.1	32	4	Aam38254 Peptide #
13	45	39.1	32	4	Aam37929 Peptide #
14	45	39.1	32	4	Abb26897 Protein #
15	45	39.1	32	4	Abb27163 Protein #
16	45	39.1	32	4	Aam77713 Human bon
17	45	39.1	32	4	Abb43988 Peptide #
18	45	39.1	32	4	Abb43988 Peptide #
19	45	39.1	32	5	Abg47050 Human pep
20	45	39.1	32	5	Abg46741 Human pep
21	45	39.1	155	4	Aau30945 Novel hum
22	45	39.1	341	7	Adf74142 Human nov
23	45	39.1	1118	7	Abg75594 Pseudomon
24	44	38.3	70	4	Aao10516 Human pol
25	44	38.3	131	3	Aab54175 Human pan

26	44	38.3	133	5	ADJ33758	Adj33758 Protein o
27	44	38.3	194	7	ADF78159	Adf78159 Human ext
28	44	38.3	243	5	AAE31060	Aae31060 Ehrlichia
29	44	38.3	413	5	AAO15858	Aao15858 Human glu
30	44	38.3	1094	5	AAU72900	Aau72900 Human met
31	44	38.3	26926	4	AAU05396	Aau05396 Human tit
32	44	38.3	26926	8	ADQ17316	Adq17316 Human sof
33	44	38.3	31267	6	ABG74786	Abg74786 Human RGS
34	43.5	37.8	884	4	ABG22489	Abg22489 Novel hum
35	43.5	37.8	884	4	ABG30084	Abg30084 Novel hum
36	43	37.4	57	8	ABO54483	AbO54483 Human gen
37	43	37.4	62	4	AAU60604	Aau60604 Propionib
38	43	37.4	62	6	ABM57123	Abm57123 Propionib
39	43	37.4	311	7	ADM26562	Adm26562 Hyperther
40	43	37.4	441	4	AAU20410	Aau20410 Human sec
41	43	37.4	451	4	ABB67753	Abb67753 Drosophil
42	43	37.4	1422	3	AAB09041	Aab09041 Hepatitis
43	42.5	37.0	261	4	AAU40132	Aau40132 Propionib
44	42.5	37.0	261	6	ABM36651	Abm36651 Propionib
45	42.5	37.0	622	6	ABU49786	Abu49786 Protein e

ALIGNMENTS

RESULT 1

AAO09931

ID AAO09931 standard; protein; 112 AA.

XX AAO09931;

XX AAO09931;

DT 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 23823.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI89862.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing

PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 23823; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and

CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 112 AA;

SQ Query Match 43.5%; Score 50; DB 4; Length 112;
 Best Local Similarity 56.2%; Pred. No. 6.2;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 PRGAPMWLRVCVCOML 16
 ||||| :||| :||
 Db 8 PRGAAIVRCVSALLE 23

RESULT 2
 AAO09705
 ID AAO09705 standard; protein; 78 AA.

XX AAO09705;
 DT 06-NOV-2001 (first entry)
 DE Human polypeptide SEQ ID NO 23597.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI89636.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 23597; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 78 AA;

SQ Query Match 42.6%; Score 49; DB 4; Length 78;
 Best Local Similarity 60.0%; Pred. NO. 6.3;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PRGAPMWLRVCVCOML 15
 ||||| :||| :||
 Db 8 PRGAPACLRCLLSALL 22

RESULT 3
 AAP91653
 ID AAP91653 standard; protein; 773 AA.

XX AAP91653;
 XX 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 22-JAN-1990 (first entry)

XX Membrane-bound aldehyde dehydrogenase.

XX Acetobacter; aldehyde dehydrogenase; acetic acid fermentation.

XX Acetobacter aceti.

XX EP332120-A.

XX 13-SEP-1989.

XX 06-MAR-1989; 89EP-00103936.

XX 08-MAR-1989; 88JP-00052709.

XX 15-FEB-1989; 89JP-00033775.

XX (NAKA-) NAKANO VINEGAR CO LTD.

XX Fukaya M, Tayama K, Tamaki T, Izumo H, Okumura H, Kawamura Y;

XX WPI; 1989-265358/37.

XX N-PSDB; AAN90816.

XX Gene for membrane-bound aldehyde dehydrogenase - derived from Acetobacter
 PT and used to transform acetic acid bacteria to enhance prodn.

XX Claim 5; Fig 3; 17pp; English.

XX This oxidises aldehyde to corresp. acid with the aid of pyrrolo-
 CC quinoline as prosthetic gp. It is used in oxidn. of ethanol to acetic
 CC acid, and in oxidative decomposition of aldehydes that give food a bad
 CC taste. The new plasmids claimed increase this enzyme content in cells. A.
 CC altoacetigenes strain MH-24 FERM BP-491 is the source DNA. (Updated on 25
 CC -MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-
 CC 2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 773 AA;

SQ Query Match 40.9%; Score 47; DB 1; Length 773;
 Best Local Similarity 61.5%; Pred. No. 1.2e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 RGAPMWLRVCVCOM 14
 ||||| :||| :||
 Db 631 RGMPWDRCAVQV 643

RESULT 4
 AAU31146
 ID AAU31146 standard; protein; 88 AA.

XX AAU31146;

XX 18-DEC-2001 (first entry)

XX Novel human secreted protein #1637.

XX Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 XX WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 XX 16-APR-2001; 2001WO-US008656.
 PF
 XX 18-APR-2000; 2000US-00552929.
 PR
 XX 26-JAN-2001; 2001US-00770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-611725/70.
 XX
 DR Nucleic acids encoding a range of human polypeptides, useful in genetic
 XX vaccination, testing and therapy.
 PT
 XX Claim 20; Page 410; 765pp; English.
 XX
 PS The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 88 AA;

 Query Match 40.0%; Score 46; DB 4; Length 88;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 10; Conservative 0; Mismatches 6; Indels 4; Gaps 1;

 QY 1 PRGAPMWLRVCVQMLTKFL 20
 Db 33 PHGPP-----CSCDFLETHFL 48

 RESULT 5
 ADC87493
 ID ADC87493 standard; protein; 274 AA.
 AC
 XX ADC87493;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human GPCR protein SEQ ID NO:1946.
 XX
 KW human; GPCR; guanosine triphosphate-binding protein coupled receptor;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 XX EP1270724-A2.
 PN
 XX 02-JAN-2003.
 PD
 XX

PF 18-JUN-2002; 2002EP-00013517.
 XX
 PR 18-JUN-2001; 2001JP-00246789.
 XX
 PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
 XX
 PI Suwa M, Asai K, Akiyama Y, Aburatani H;
 XX WPI; 2003-315783/31.
 DR N-PSDB; ADC87492.
 XX
 PT New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.
 XX
 PS Claim 2; SEQ ID NO 1946; 28pp; English.
 XX
 CC The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The protein
 CC sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.
 XX
 SQ Sequence 274 AA;

 Query Match 40.0%; Score 46; DB 7; Length 274;
 Best Local Similarity 53.3%; Pred. No. 63;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

 QY 1 PRGAPMWLRVCVQML 15
 Db 70 PCGLPLDFSRVCVFL 84

 RESULT 6
 ARG01589
 ID ARG01589 standard; protein; 634 AA.
 XX
 AC ARG01589;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #1580.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200175067-A2.
 PN
 XX 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR
 XX 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 XX N-PSDB; AAS65776.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PT
 XX

PS Claim 20; SEQ ID NO 31948; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (II) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders CC involving aberrant protein expression or biological activity. The CC polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic CC amino acid sequences of the invention. Note: The sequence data for this CC patent did not appear in the printed specification, but was obtained in CC electronic format directly from WIPO at CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 634 AA;

SQ

Query Match 40.0%; Score 46; DB 4; Length 634;

Best Local Similarity 63.6%; Pred. No. 1.4e+02;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GAPMWRVCVCQ 13

DB 423 GASSWRRLCQC 433

RESULT 7

ADE95598

ID ADE95598 standard; protein; 195 AA.

AC ADE95598;

XX

DT 12-FEB-2004 (first entry)

XX

DE Human NOVX22c protein.

XX NOVX protein; biochemical stimulation; physiological stimulation; CC cardiant; antiarteriosclerotic; hypotensive; cytostatic; anorectic; CC antirheumatic; antiarthritic; antidiabetic; nephrotropic; dermatological; CC immunosuppressive; anti-HIV; antiinflammatory; neuroprotective; CC neutropic; antipsoriatic; antiparkinsonian; antiasthmatic; neuroleptic; CC antidepressant; antiallergic; gynaecological; gene therapy; vaccine; CC NOVX-associated disorder; cardiomyopathy; atherosclerosis; hypertension; CC cancer; obesity; rheumatoid arthritis; diabetes; glomerulonephritis; CC psoriasis; skin disorder; AIDS; inflammation; multiple sclerosis; CC Alzheimer's disease; Parkinson's disease; asthma; schizophrenia; CC depression; allergy; fertility disorder; NOVX22c.

XX Homo sapiens.

OS

XX

PN WO2003050245-A2.

XX

XX 19-JUN-2003.

XX

PF 03-DEC-2002; 2002WO-US038594.

XX

XX 05-DEC-2001; 2001US-0336600P.

PR 07-DEC-2001; 2001US-0338285P.

PR 12-DEC-2001; 2001US-0341346P.

PR 17-DEC-2001; 2001US-0341477P.

PR 17-DEC-2001; 2001US-0341540P.

PR 20-DEC-2001; 2001US-0342592P.

PR 27-DEC-2001; 2001US-0344297P.

PR 31-DEC-2001; 2001US-0344903P.

PR 17-APR-2002; 2002US-0373288P.

PR 15-MAY-2002; 2002US-0380981P.

PR 17-MAY-2002; 2002US-0381495P.

PR 28-MAY-2002; 2002US-0383534P.

PR 28-MAY-2002; 2002US-0383744P.

PR 29-MAY-2002; 2002US-0383829P.

PR 29-MAY-2002; 2002US-0384024P.

PR 07-AUG-2002; 2002US-0401788P.

PR 26-AUG-2002; 2002US-0406353P.

PR 31-OCT-2002; 2002US-00401788.

PR 02-DEC-2002; 2002US-00406353.

XX (CURA-) CURAGEN CORP.

XX

XX Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Chillakuru RA, Edinger SR, Gerlach VL, Gorman L, Gould-Rothberg BE, Guo X; Jeffers ME, Ji W, Li L, Malyankar UM, Miller CE, Murphy R; Patturajan M, Payman JA, Rastelli L, Rieger DK, Shenoy SG; Smithson G, Starling G, Taupier RJ, Voss EZ, Zhong H, Zhong M; WPI; 2003-513974/48.

DR N-PSDB; ADE95597.

XX

XX New NOVX polypeptides and nucleic acids, useful for preventing or PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy, PT atherosclerosis or diabetes, and in chromosome mapping, tissue typing or PT pharmacogenomics.

XX

PS Claim 2; SEQ ID NO 130; 211pp; English.

XX

CC This invention relates to novel NOVX proteins, and the DNA sequence which CC encode them, having properties related to stimulation of biochemical or CC physiological responses in a cell, a tissue, an organ or an organism. CC Compounds which modulate the proteins of the invention may have cardiant, CC antiarteriosclerotic, hypotensive, cytostatic, anorectic, antirheumatic, CC antiarthritic, antidiabetic, nephrotropic, dermatological, CC immunosuppressive, anti-HIV, antiinflammatory, neuroprotective, CC neutropic, antipsoriatic, antiparkinsonian, antiasthmatic, neuroleptic, CC antidepressant, antiallergic or gynaecological activities. The DNA CC sequences of the invention may be useful for gene therapy whilst the CC protein sequences may allow the development of a vaccine. The protein is CC useful in the manufacture of a medicament for treating a syndrome CC associated with a human disease. The invention may be useful in CC diagnosing, treating or preventing NOVX-associated disorders, for example CC cardiomyopathy, atherosclerosis, hypertension, cancer, obesity, CC rheumatoid arthritis, diabetes, glomerulonephritis, psoriasis, skin CC disorders, AIDS, inflammation, multiple sclerosis, Alzheimer's disease, CC Parkinson's disease, asthma, schizophrenia, depression, allergies or CC fertility disorders. The nucleic acids may further be used as CC hybridisation probes, in chromosome mapping, tissue typing, preventive CC medicine, and pharmacogenomics. The present sequence is the amino acid CC sequence of the human NOVX22c protein of the invention.

XX

SQ Sequence 195 AA;

Query Match 39.6%; Score 45.5; DB 7; Length 195;

Best Local Similarity 66.7%; Pred. No. 55;

Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 PRGAPMWRVCVC 12

DB 184 PLGAPQWV-CVC 194

RESULT 8

ADE95594

ID ADE95594 standard; protein; 264 AA.

XX

XX ADE95594;

XX

DT 12-FEB-2004 (first entry)

XX

DE Human NOVX22a protein.

CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have application in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
CC SQ Sequence 482 AA;

Query Match 39.6%; Score 45.5; DB 4; Length 482;
Best Local Similarity 42.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 5; Gaps 1;

QY 1 PRGAPMWL-----RCVCOM 14
||| |::|
DB 244 PRTHPVVWVTRGRCRCTQL 262

RESULT 10
AAM21627
ID AAM21627 standard; protein; 32 AA.
XX AC AAM21627;
XX DT 12-OCT-2001 (first entry)
XX DE Peptide #8061 encoded by probe for measuring cervical gene expression.
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer.
XX OS Homo sapiens.
XX XN WO200157278-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000670.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PF WPI; 2001-488901/53.
XX DR WPI; 2001-488901/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human cervical epithelial cells.
XX PS Claim 27; SEQ ID NO 26453; 487pp; English.
XX CC The present invention relates to human single exon nucleic acid probes
XX CC (SENP: see AA110668-AA128459). The present sequence is a peptide encoded
XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX CC can be used to produce a single exon microarray, which can be used for
XX CC measuring human gene expression in a sample derived from human cervical

CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
CC
CC SQ Sequence 32 AA;

Query Match 39.1%; Score 45; DB 4; Length 32;
Best Local Similarity 85.7%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 MWLRVCV 12
||:||||
DB 26 MMWRCVC 32

RESULT 11
ABB43988
ID ABB43988 standard; peptide; 32 AA.
XX AC ABB43988;
XX DT 04-FEB-2002 (first entry)
XX DE Peptide #11494 encoded by human foetal liver single exon probe.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX OS Homo sapiens.
XX XN WO200157277-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000659.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PF WPI; 2001-483447/52.
XX DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human fetal liver.
XX PS Claim 27; SEQ ID NO 36623; 639pp + Sequence Listing; English.
XX CC The invention relates to a single exon nucleic acid probe for measuring
XX CC human gene expression in a sample derived from human foetal liver. The
XX CC single exon nucleic acid probes may be used for predicting, measuring and
XX CC displaying gene expression in samples derived from human fetal liver. The
XX CC present sequence is a peptide encoded by a single exon nucleic acid probe
XX CC of the invention. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 32 AA;

Query Match 39.1%; Score 45; DB 4; Length 32;
Best Local Similarity 85.7%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 MWLRVCV 12
||:||||


```

Db      26 MWLRVCV 32
RESULT 12
AAM38254
ID AAM38254 standard; protein; 32 AA.
XX
AC AAM38254;
XX
XX 17-OCT-2001 (first entry)
XX
DE Peptide #12291 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
XX Claim 27; SEQ ID NO 38523; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see AAI31315-AAI57546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders
XX
XX SQ Sequence 32 AA;
XX
XX Query Match 39.1%; Score 45; DB 4; Length 32;
XX Best Local Similarity 85.7%; Pred. No. 12;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 6 MWLRVCV 12
DB 26 MWLRVCV 32
||:||||
||:||||
RESULT 14
ABB26897
ID ABB26897 standard; protein; 32 AA.
XX
XX ABB26897;
XX
XX 23-JAN-2002 (first entry)
XX
XX Protein #8896 encoded by probe for measuring heart cell gene expression.
XX
XX Human; gene expression; heart; microarray; vascular system;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease.
XX
XX OS Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
XX Claim 27; SEQ ID NO 38523; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see AAI31315-AAI57546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders
XX
XX SQ Sequence 32 AA;
XX
XX Query Match 39.1%; Score 45; DB 4; Length 32;
XX Best Local Similarity 85.7%; Pred. No. 12;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 6 MWLRVCV 12
DB 26 MWLRVCV 32
||:||||
||:||||
RESULT 13
AAM37929
ID AAM37929 standard; protein; 32 AA.
XX
XX AAM37929;
XX
XX 17-OCT-2001 (first entry)
XX
XX Peptide #11966 encoded by probe for measuring placental gene expression.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.

```

```

PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 28667; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 32 AA;
Query Match 39.1%; Score 45; DB 4; Length 32;
Best Local Similarity 85.7%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 6 MWLRVCV 12
Db 26 MWMRCVC 32
||:||||
||:||||

RESULT 15
ABB27163
ID ABB27163 standard; protein; 32 AA.
XX
AC ABB27163;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #9162 encoded by probe for measuring heart cell gene expression.
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.

```

```

XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 28933; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 32 AA;
Query Match 39.1%; Score 45; DB 4; Length 32;
Best Local Similarity 85.7%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 6 MWLRVCV 12
Db 26 MWMRCVC 32
||:||||
||:||||

Search completed: October 26, 2004, 06:45:14
Job time : 86.5 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:45:25 ; Search time 64.25 Seconds
(without alignments)
100.781 Million cell updates/sec

Title: US-10-066-965A-4

Perfect score: 115

Sequence: 1 PRGAPWMLRCVCOMLETKFL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	115	100.0	20	14	US-10-066-965A-4
2	115	100.0	20	14	US-10-066-965A-23
3	96	83.5	20	14	US-10-066-965A-3
4	96	83.5	20	14	US-10-066-965A-22
5	50	43.5	1044	14	US-10-369-493-22713
6	49	42.6	103	16	US-10-437-963-142466
7	46	40.0	274	14	US-10-017-161-2300
8	46	40.0	274	14	US-10-232-798-1946
9	45.5	39.6	195	15	US-10-309-290-130
10	45.5	39.6	264	15	US-10-309-290-126
11	45	39.1	32	9	US-09-864-761-42195
12	45	39.1	32	9	US-09-864-761-42461
13	45	39.1	260	14	US-10-156-761-10442

14	45	39.1	294	16	US-10-437-963-133679	Sequence 133679, Ap
15	45	39.1	469	14	US-10-369-493-3310	Sequence 3310, Ap
16	44	38.3	88	16	US-10-767-701-55432	Sequence 55432, A
17	44	38.3	131	9	US-09-925-297-627	Sequence 627, App
18	44	38.3	243	14	US-10-081-051-26	Sequence 26, Appli
19	44	38.3	413	13	US-10-106-534-6	Sequence 6, Appli
20	44	38.3	1094	15	US-10-275-107-60	Sequence 60, Appli
21	44	38.3	26926	9	US-09-759-508B-2	Sequence 2, Appli
22	43.5	37.8	371	15	US-10-425-114-42489	Sequence 42489, A
23	43.5	37.8	526	15	US-10-424-599-239333	Sequence 239333, A
24	43	37.4	57	14	US-10-029-386-28117	Sequence 28117, A
25	43	37.4	98	16	US-10-437-963-125793	Sequence 125793, A
26	43	37.4	128	16	US-10-437-963-125130	Sequence 125130, A
27	43	37.4	163	15	US-10-424-599-235647	Sequence 235647, A
28	43	37.4	536	16	US-10-437-963-123986	Sequence 123986, A
29	43	37.4	575	16	US-10-437-963-121604	Sequence 121604, A
30	43	37.4	645	14	US-10-156-761-7854	Sequence 7854, Ap
31	43	37.4	1422	8	US-08-424-550B-86	Sequence 86, Appli
32	42.5	37.0	51	15	US-10-424-599-267231	Sequence 267231, A
33	42.5	37.0	73	15	US-10-424-599-144792	Sequence 144792, A
34	42.5	37.0	98	16	US-10-437-963-105197	Sequence 105197, A
35	42.5	37.0	325	16	US-10-437-963-166717	Sequence 166717, A
36	42.5	37.0	479	16	US-10-437-963-191710	Sequence 191710, A
37	42.5	37.0	489	16	US-10-437-963-197107	Sequence 197107, A
38	42.5	37.0	622	15	US-10-282-122A-77710	Sequence 77710, A
39	42	36.5	56	13	US-10-001-843-183	Sequence 183, App
40	42	36.5	56	15	US-10-424-599-236745	Sequence 236745, A
41	42	36.5	73	15	US-10-424-599-280804	Sequence 280804, A
42	42	36.5	86	15	US-10-424-599-211988	Sequence 211988, A
43	42	36.5	104	15	US-10-424-599-144556	Sequence 144556, A
44	42	36.5	133	14	US-10-106-698-6516	Sequence 6516, Ap
45	42	36.5	173	15	US-10-425-114-60423	Sequence 60423, A

ALIGNMENTS

RESULT 1
US-10-066-965A-4
; Sequence 4, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYPT 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066, 965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-4

Query Match 100.0%; Score 115; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRGAPWMLRCVCOMLETKFL 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 PRGAPWMLRCVCOMLETKFL 20

RESULT 2
US-10-066-965A-23
; Sequence 23, Application US/10066965A
; Publication No. US20030143626A1

```

; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYP 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-23

Query Match      100.0%; Score 115; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PRGAPMRLRCVQCMLETKFL 20
Db      1 PRGAPMRLRCVQCMLETKFL 20

RESULT 3
US-10-066-965A-3
; Sequence 3, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYP 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-3

Query Match      83.5%; Score 96; DB 14; Length 20;
Best Local Similarity 85.0%; Pred. No. 3.8e-07;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 PRGAPMRLRCVQCMLETKFL 20
Db      1 PRGAPMRLRCVQCMLETKFL 20

RESULT 4
US-10-066-965A-22
; Sequence 22, Application US/10066965A
; Publication No. US20030143626A1
; GENERAL INFORMATION:
; APPLICANT: COLAS, PIERRE
; APPLICANT: BRENT, ROGER
; APPLICANT: COHEN, BARAK A.
; TITLE OF INVENTION: TARGETED MODIFICATION OF INTRACELLULAR COMPOUNDS
; FILE REFERENCE: EGYP 3.0-015
; CURRENT APPLICATION NUMBER: US/10/066,965A
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 33

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-066-965A-22

Query Match      83.5%; Score 96; DB 14; Length 20;
Best Local Similarity 85.0%; Pred. No. 3.8e-07;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 PRGAPMRLRCVQCMLETKFL 20
Db      1 PRGAPMRLRCVQCMLETKFL 20

RESULT 5
US-10-369-493-22713
; Sequence 22713, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22713
; LENGTH: 1044
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-22713

Query Match      43.5%; Score 50; DB 14; Length 1044;
Best Local Similarity 53.3%; Pred. No. 83;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      6 MWLRCVQCMLETKFL 20
Db      52 VVYRCIEWLLETRFL 66

RESULT 6
US-10-437-963-142466
; Sequence 142466, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142466
; LENGTH: 103

```

```
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4346C.1.pep
US-10-437-963-142466

Query Match      42.6%; Score 49; DB 16; Length 103;
Best Local Similarity 45.0%; Pred. No. 14;
Matches 9; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Oy 1 PRGAPMWLRVCVQMLLETKFL 20
Db 17 PRARPCWLSCSCAAALLAFL 36

RESULT 7
US-10-161-161-2300
; Sequence 2300, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2300
; TYPE: PRT
; LENGTH: 274
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (25)..(25)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (45)..(45)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (51)..(51)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (187)..(187)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (222)..(222)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (248)..(248)
; OTHER INFORMATION: Variable amino acid
US-10-017-161-2300

Query Match      40.0%; Score 46; DB 14; Length 274;
Best Local Similarity 53.3%; Pred. No. 96;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 1 PRGAPMWLRVCVQML 15
Db 70 PCGLPLFSRCVCVFL 84

RESULT 8
US-10-292-798-1946
; Sequence 1946, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1946
; TYPE: PRT
; LENGTH: 274
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (25)..(25)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (45)..(45)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (51)..(51)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (187)..(187)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (222)..(222)
; OTHER INFORMATION: Variable amino acid
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (248)..(248)
; OTHER INFORMATION: Variable amino acid
US-10-292-798-1946

Query Match      40.0%; Score 46; DB 14; Length 274;
Best Local Similarity 53.3%; Pred. No. 96;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 1 PRGAPMWLRVCVQML 15
Db 70 PCGLPLFSRCVCVFL 84

RESULT 9
US-10-309-290-130
; Sequence 130, Application US/10309290
; Publication No. US20040023241A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Chilikuru, Rajeev A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Gorman, Linda
; APPLICANT: Gould-Rothberg, Bonnie E.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Jeffers, Michael E.
; APPLICANT: Ji, Weizhen
```

```

; APPLICANT: Li, Li
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Miller, Charles E.
; APPLICANT: Murphey, Ryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Smithson, Glenda
; APPLICANT: Starling, Gary
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-502A
; CURRENT APPLICATION NUMBER: US/10/309,290
; PRIOR FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/336,600
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,285
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/341,346
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,297
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/344,903
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/373,288
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/380,981
; PRIOR FILING DATE: 2002-05-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 274
; SOFTWARE: CuraSequist version 0.1
; SEQ ID NO 130
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-309-290-130

```

```

Query Match      39.6%; Score 45.5; DB 15; Length 195;
Best Local Similarity 66.7%; Pred. No. 84;
Matches      8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

```

```

Qy      1 PRGAPMMLRCVC 12
      |||||: |||
Db      184 PLGAPQWV-CVC 194

```

```

RESULT 10
US-10-309-290-126
; Sequence 126, Application US/10309290
; Publication No. US20040023241A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Chillakuru, Rajeev A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Gorman, Linda
; APPLICANT: Gould-Rothberg, Bonnie E.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Jeffers, Michael E.

```

```

; APPLICANT: Ji, Weizhen
; APPLICANT: Li, Li
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Miller, Charles E.
; APPLICANT: Murphey, Ryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Smithson, Glenda
; APPLICANT: Starling, Gary
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-502A
; CURRENT APPLICATION NUMBER: US/10/309,290
; PRIOR FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/336,600
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,285
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/341,346
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,297
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/344,903
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/373,288
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/380,981
; PRIOR FILING DATE: 2002-05-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 274
; SOFTWARE: CuraSequist version 0.1
; SEQ ID NO 126
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-309-290-126

```

```

Query Match      39.6%; Score 45.5; DB 15; Length 264;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches      8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

```

```

Qy      1 PRGAPMMLRCVC 12
      |||||: |||
Db      253 PLGAPQWV-CVC 263

```

```

RESULT 11
US-09-864-761-42195
; Sequence 42195, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312

```

```

PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 42195
LENGTH: 32
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC021876.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.8
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 10
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
OTHER INFORMATION: EST_HUMAN HIT: C03546.1, EVALUATE 3.00e-13

```

RESULT 13
US-10-156-761-10442
; Sequence 10442, Application US/10156761
; Publication No. US20030119018A1

RESULT 12
US-09-864-761-42461
; Sequence 42461, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; APPLICANT: TWENTY-ONE HEMAN, GENOME-DETERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

```
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10442
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10442

Query Match          39.1%; Score 45; DB 14; Length 260;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 PRGAPMWLRVCV 12
    |||||
Db 68 PRGAPRWPAVC 79

RESULT 14
US-10-437-963-133679
; Sequence 133679, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133679
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_35525C.1.pep
US-10-437-963-133679

Query Match          39.1%; Score 45; DB 16; Length 294;
Best Local Similarity 47.4%; Pred. No. 1.4e+02;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 RGAPMWLRVCVCMLETKFL 20
    |||||
Db 211 RGAPVALACOCALLGCNLL 229

RESULT 15
US-10-369-493-3310
; Sequence 3310, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
```

```
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3310
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3310

Query Match          39.1%; Score 45; DB 14; Length 469;
Best Local Similarity 42.1%; Pred. No. 2.2e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 RGAPMWLRVCVCMLETKFL 20
    |||||
Db 258 RGAEWVRDVCPSLDGKYI 276

Search completed: October 26, 2004, 07:10:19
Job time : 64.25 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:20:00 ; Search time 22 Seconds
(without alignments)
60.289 Million cell updates/sec

Title: US-10-066-965A-4

Perfect score: 115
Sequence: 1 PRGAPMWLRVCQMLETKFL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	39.1	1118	4	US-09-252-991A-24340
2	43.5	37.8	77	4	Sequence 24340, A
3	43	37.4	154	4	Sequence 4938, Ap
4	43	37.4	239	4	Sequence 59364, A
5	43	37.4	447	3	Sequence 43959, A
6	43	37.4	1422	4	Sequence 7, Appl
7	43	37.4	1422	4	Sequence 86, Appl
8	43	37.4	1422	4	Sequence 86, Appl
9	43	37.4	1422	4	Sequence 86, Appl
10	42	36.5	88	3	Sequence 5158, Ap
11	42	36.5	91	4	Sequence 38687, A
12	42	36.5	91	4	Sequence 53884, A
13	42	36.5	336	4	Sequence 22897, A
14	41	35.7	70	4	Sequence 7736, Ap
15	41	35.7	113	4	Sequence 39796, A
16	41	35.7	113	4	Sequence 55013, A
17	41	35.7	175	4	Sequence 31113, A
18	41	35.7	233	4	Sequence 12151, A
19	41	35.7	359	4	Sequence 36276, A
20	41	35.7	359	4	Sequence 51493, A
21	41	35.7	992	1	Sequence 2, Appl
22	41	35.7	992	1	Sequence 2, Appl
23	41	35.7	992	1	Sequence 2, Appl
24	41	35.7	992	1	Sequence 2, Appl
25	41	35.7	992	1	Sequence 2, Appl
26	41	35.7	992	1	Sequence 2, Appl
27	41	35.7	992	2	Sequence 2, Appl

28	41	35.7	992	4	US-09-872-136B-2	Sequence 2, Appl
29	41	35.7	992	5	PCT-US92-02750-2	Sequence 2, Appl
30	41	35.7	992	5	PCT-US92-05401-2	Sequence 2, Appl
31	41	35.7	992	5	PCT-US92-09893-2	Sequence 4, Appl
32	41	35.7	993	1	US-07-977-451-4	Sequence 4, Appl
33	41	35.7	993	1	US-08-252-517-4	Sequence 2, Appl
34	41	35.7	993	1	US-08-183-211-2	Sequence 4, Appl
35	41	35.7	993	1	US-07-906-397A-4	Sequence 4, Appl
36	41	35.7	993	1	US-08-222-299-4	Sequence 4, Appl
37	41	35.7	993	2	US-09-021-324-4	Sequence 4, Appl
38	41	35.7	993	2	US-08-434-878-4	Sequence 4, Appl
39	41	35.7	993	4	US-09-872-136B-4	Sequence 4, Appl
40	41	35.7	993	5	PCT-US92-09893-4	Sequence 2, Appl
42	41	35.7	993	5	PCT-US95-00176A-2	Sequence 4, Appl
43	41	35.7	1000	1	PCT-US95-03718-4	Sequence 2, Appl
44	41	35.7	1000	2	US-08-222-299-2	Sequence 2, Appl
45	41	35.7	1000	2	US-08-434-878-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-252-991A-24340
; Sequence 24340, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24340
; LENGTH: 1118
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24340

Query Match 39.1%; Score 45; DB 4; Length 1118;
Best Local Similarity 77.8%; Pred. No. 74;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PRGAPMWLR 9
DB 489 PRGSPAWLR 497

RESULT 2
US-09-621-976-4938
; Sequence 4938, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4938
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

RESULT 5

Query Match 37.4%; Score 43; DB 4; Length 1422;
Best Local Similarity 36.0%; Pred. No. 2e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 6; Gaps 1;

QY 1 PRGAP-----MWLRVCQMLETKF 19
Db 1262 PHXSPACKVVIWLTCTRCQWGLKF 1286

RESULT 7

US-08-488-446-86
; Sequence 86, Application US/08488446
; Patent No. 6558998
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1422 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-488-446-86

Query Match 37.4%; Score 43; DB 4; Length 1422;
Best Local Similarity 36.0%; Pred. No. 2e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 6; Gaps 1;

QY 1 PRGAP-----MWLRVCQMLETKF 19
Db 1262 PHXSPACKVVIWLTCTRCQWGLKF 1286

RESULT 8
US-08-467-344A-86

;; TITLE: Encoding a-Galactosidase
;; JOURNAL: Molecular and General Genetics
;; VOLUME: 233
;; PAGES: 404-410
;; DATE: 1992
;; DOCUMENT NUMBER:
;; FILING DATE:
;; PUBLICATION DATE:
;; RELEVANT RESIDUES IN SEQ ID NO:
US-09-070-356-7

Query Match 37.4%; Score 43; DB 3; Length 447;
Best Local Similarity 42.1%; Pred. No. 60;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 PRGAPMWLRVCQMLETKF 19
Db 43 PMGPNWAWEMCDLNETLF 61

RESULT 6

US-08-469-260A-86
; Sequence 86, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,260A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1422 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-260A-86

; Sequence 86, Application US/08467344A
; Patent No. 6986568
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS

; TAMI J. PILOT-MATIAS
; GEORGE J. DAWSON
; GEORGE G. SCHLAUDER
; SURESH M. DESAI
; THOMAS P. LEARY
; ANTHONY SCOTT MUERHOFF
; JAMES C. ERKER
; SHERI L. BULJK
; ISA K. MUSHAWAR

; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; REAGENTS AND METHODS FOR THEIR USE

; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL

; COUNTRY: USA
; ZIP: 60064-3500

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; FILING DATE: 07-Jun-1995

; APPLICATION NUMBER: US/08/467,344A

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/424,550

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: FOREMSKI, PRISCILLA E.

; REGISTRATION NUMBER: 33,207

; REFERENCE/DOCKET NUMBER: 5527.PC.01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 708-937-6365

; TELEFAX: 708-938-2623

; INFORMATION FOR SEQ ID NO: 86:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1422 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 86:

US-08-467-344A-86

Query Match 37.4%; Score 43; DB 4; Length 1422;
Best Local Similarity 36.0%; Pred. No. 2e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 6; Gaps 1;

Qy 1 PRGAP-----MWLRVCVCMLETKF 19
Db 1262 PHXSPACKKVVILTCRCQWGLKF 1286

RESULT 9

US-08-424-550B-86

; Sequence 86, Application US/08424550B

; Patent No. 6720166

; GENERAL INFORMATION:

; APPLICANT: JOHN N. SIMONS

; TAMI J. PILOT-MATIAS

; GEORGE J. DAWSON

; GEORGE G. SCHLAUDER

; SURESH M. DESAI

; THOMAS P. LEARY

; ANTHONY SCOTT MUERHOFF

; APPLICANT: ANTHONY SCOTT MUERHOFF

; APPLICANT: ANTHONY SCOTT MUERHOFF

; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BULJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D

; STREET: 100 ABBOTT PARK ROAD

; CITY: ABBOTT PARK

; STATE: IL

; COUNTRY: USA

; ZIP: 60064-3500

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/424,550B

; FILING DATE:

; CLASSIFICATION: 435435

; ATTORNEY/AGENT INFORMATION:

; NAME: FOREMSKI, PRISCILLA E.

; REGISTRATION NUMBER: 33,207

; REFERENCE/DOCKET NUMBER: 5527.PC.01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 708-937-6365

; TELEFAX: 708-938-2623

; INFORMATION FOR SEQ ID NO: 86:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1422 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-424-550B-86

Query Match 37.4%; Score 43; DB 4; Length 1422;
Best Local Similarity 36.0%; Pred. No. 2e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 6; Gaps 1;

Qy 1 PRGAP-----MWLRVCVCMLETKF 19
Db 1262 PHXSPACKKVVILTCRCQWGLKF 1286

RESULT 10

US-09-134-001C-5158

; Sequence 5158, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; REAGENTS AND METHODS FOR THEIR USE

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 5158

; LENGTH: 88

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5158

Query Match 36.5%; Score 42; DB 3; Length 88;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
QY      6 MWLRVCQML 15
      |||: |||: |||
Db      26 MWINCCVCL 35

RESULT 11
US-09-270-767-38667
; Sequence 38667, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38667
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-38667

Query Match      36.5%; Score 42; DB 4; Length 91;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAPMWL 8
      |||||: |||
Db      65 PRGAPVWV 72

RESULT 12
US-09-270-767-53884
; Sequence 53884, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53884
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-53884

Query Match      36.5%; Score 42; DB 4; Length 91;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRGAPMWL 8
      |||||: |||
Db      65 PRGAPVWV 72

RESULT 13
US-09-252-991A-22897
; Sequence 22897, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
```

```
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22897
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22897

Query Match      36.5%; Score 42; DB 4; Length 336;
Best Local Similarity 63.6%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 PRGAPMWLRCV 11
      |||: |||: |||
Db      248 PRGLEEWARCV 258

RESULT 14
US-09-621-976-7736
; Sequence 7736, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7736
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7736

Query Match      35.7%; Score 41; DB 4; Length 70;
Best Local Similarity 45.5%; Pred. No. 18;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      5 PMWLRCVCQML 15
      |||: |||: |||
Db      45 PLWSRCIVEIL 55

RESULT 15
US-09-270-767-39796
; Sequence 39796, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39796
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39796

Query Match      35.7%; Score 41; DB 4; Length 113;
Best Local Similarity 63.6%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 RGAPMWLRCVC 12
      |||: |||: |||
Db      101 RGRGWLHCVC 111
```

Tue Oct 26 08:02:26 2004

us-10-066-965a-4.rai

Page 6

Search completed: October 26, 2004, 06:47:58
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 07:05:46 ; Search time 16.5 Seconds
(without alignments)
116.626 Million cell updates/sec

Title: US-10-066-965A-1

Perfect score: 20

Sequence: 1 QVSLWALGWRRLRRYGNW 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

PIR 79: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	40.0	297	C83311	hypothetical prote
2	8	40.0	2	AG2839	hypothetical prote
3	8	40.0	518	A97617	probable transport
4	7	35.0	571	C84356	hypothetical prote
5	6	30.0	174	T46382	hypothetical prote
6	6	30.0	191	F82725	hypothetical prote
7	6	30.0	266	D83210	hypothetical prote
8	6	30.0	331	A83377	conserved hypotet
9	6	30.0	361	JN0530	thioesterase homol
10	6	30.0	379	S67856	GumG protein - Xan
11	6	30.0	411	F87644	transcription regu
12	6	30.0	426	C83212	rhamnosyltransfera
13	6	30.0	426	B53652	rhamnosyltransfera
14	6	30.0	435	S55273	amine oxidase (fla
15	6	30.0	522	B86465	protein F319.15 [
16	6	30.0	560	S46096	probable membrane
17	6	30.0	636	S76730	hypothetical prote
18	6	30.0	880	T02245	hypothetical prote
19	6	30.0	1758	F88559	protein C4BP4.4b [
20	6	30.0	1767	S60124	transport protein
21	5	25.0	60	QXBP7L	hypothetical prote
22	5	25.0	65	D83827	Na+/H+ antiporter
23	5	25.0	75	A72733	hypothetical prote
24	5	25.0	84	S72700	probable ketoacyl
25	5	25.0	89	I56009	MHC H1A-A2-alpha-2
26	5	25.0	91	I58989	MHC H2-K transplan
27	5	25.0	91	I57611	MHC K-bm6 transpla
28	5	25.0	91	I59068	MHC class I H2-K-b
29	5	25.0	92	H30401	NADH2 dehydrogenas

30	5	25.0	92	D30401	NADH2 dehydrogenas
31	5	25.0	92	E30401	NADH2 dehydrogenas
32	5	25.0	92	E30401	NADH2 dehydrogenas
33	5	25.0	92	G30401	NADH2 dehydrogenas
34	5	25.0	97	S45490	NADH2 dehydrogenas
35	5	25.0	98	T09865	NADH2 dehydrogenas
36	5	25.0	98	T09955	NADH2 dehydrogenas
37	5	25.0	98	JQ0906	bombyxin A-3 homol
38	5	25.0	99	JQ0902	bombyxin A-1 homol
39	5	25.0	109	AG1237	B. subtilis YneA p
40	5	25.0	113	H89992	hypothetical prote
41	5	25.0	116	H75591	hypothetical prote
42	5	25.0	116	H72489	hypothetical prote
43	5	25.0	118	S51207	cruxhalorhodopsin-
44	5	25.0	118	D97046	hypothetical prote
45	5	25.0	120	B71020	hypothetical prote
46	5	25.0	122	T05357	ubiquinol-cytochro
47	5	25.0	125	T03359	gene e8 protein -
48	5	25.0	126	I63123	cytochrome-c oxida
49	5	25.0	126	I48089	cytochrome-c oxida
50	5	25.0	126	I63124	cytochrome-c oxida
51	5	25.0	126	I63121	cytochrome-c oxida
52	5	25.0	126	I63125	cytochrome-c oxida
53	5	25.0	126	I63122	cytochrome-c oxida
54	5	25.0	126	I63116	cytochrome-c oxida
55	5	25.0	126	I63117	cytochrome-c oxida
56	5	25.0	126	I63120	cytochrome-c oxida
57	5	25.0	126	I63119	cytochrome-c oxida
58	5	25.0	126	I63118	cytochrome-c oxida
59	5	25.0	126	I63128	cytochrome-c oxida
60	5	25.0	126	I63126	cytochrome-c oxida
61	5	25.0	126	I63127	cytochrome-c oxida
62	5	25.0	126	I48088	cytochrome-c oxida
63	5	25.0	126	S77183	hypothetical prote
64	5	25.0	133	F45345	trans-regulatory s
65	5	25.0	133	1 VKLJCE	trans-regulatory s
66	5	25.0	133	S57492	cytochrome-c oxida
67	5	25.0	133	S57491	cytochrome-c oxida
68	5	25.0	133	AI0999	conserved hypotet
69	5	25.0	136	D55511	bphF 3'-region hyp
70	5	25.0	137	I80172	class I histocompa
71	5	25.0	137	I80175	class I histocompa
72	5	25.0	137	I80173	class I histocompa
73	5	25.0	137	I80176	class I histocompa
74	5	25.0	137	I80174	class I histocompa
75	5	25.0	137	I38875	MHC class I antige
76	5	25.0	137	I38860	MHC class I antige
77	5	25.0	137	I38874	MHC class I antige
78	5	25.0	137	I38876	MHC class I antige
79	5	25.0	138	AI1156	hypothetical prote
80	5	25.0	138	A84482	hypothetical prote
81	5	25.0	143	S23801	pathogenesis-relat
82	5	25.0	149	AE0812	conserved hypotet
83	5	25.0	152	T44544	conserved hypotet
84	5	25.0	152	F69509	hypothetical prote
85	5	25.0	154	B83274	phosphotyrosine pr
86	5	25.0	156	AG0754	patch repair prote
87	5	25.0	157	E64008	hypothetical prote
88	5	25.0	157	E83481	hypothetical prote
89	5	25.0	163	B39384	finger protein HPF
90	5	25.0	166	I69004	histocompatibility
91	5	25.0	167	I69006	histocompatibility
92	5	25.0	167	H81252	NADH2 dehydrogenas
93	5	25.0	173	G86824	hypothetical prote
94	5	25.0	178	T08444	hypothetical prote
95	5	25.0	179	I54462	MHC RT21 protein -
96	5	25.0	180	D45831	MHC class I histoc
97	5	25.0	181	I79640	MHC cell surface a
98	5	25.0	181	I59188	MHC cell surface g
99	5	25.0	181	C75381	hypothetical prote
100	5	25.0	181	E83342	hypothetical prote

ALIGNMENTS

RESULT 1
 C83311
 hypothetical protein PA2661 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: C83311
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
 A:Reference number: AB2950; MUID:20437337; PMID:10984043
 A:Accession: C83311
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-297 <STO>
 A:Cross-references: UNIPROT:Q910H7; GB:AE004695; GB:AE004730; NID:g9948730; PIDN:AA0604
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2661

Query Match 40.0%; Score 8; DB 2; Length 297;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 LQWRWLR 15
 DB 72 LQWRWLR 79

RESULT 2
 AG2839
 hypothetical protein Atu2143 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C:Accession: AG2839
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AG2839
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-504 <KUR>
 A:Cross-references: UNIPROT:Q8UDI2; GB:AE008688; PIDN:AAL43133.1; PID:g17740607; GSPDB:C
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu2143
 A:Map position: circular chromosome

Query Match 40.0%; Score 8; DB 2; Length 504;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 VNSLWALG 9
 DB 29 VNSLWALG 36

RESULT 3
 A97617
 probable transport protein (PA4023) [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C:Accession: A97617
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: A97617
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-518 <KUR>
 A:Cross-references: UNIPROT:Q8UDI2; GB:AE007869; PIDN:AAK87890.1; PID:g15157282; GSPDB:G
 C:Genetics:
 A:Gene: AGR_C_3885
 A:Map position: circular chromosome

Query Match 40.0%; Score 8; DB 2; Length 518;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VNSLWALG 9
 DB 43 VNSLWALG 50

RESULT 4
 C84356
 hypothetical protein Vng2064h [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C:Accession: C84356
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li, L.
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: AB4160; MUID:20504483; PMID:11016950
 A:Accession: C84356
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-571 <STO>
 A:Cross-references: UNIPROT:Q9HNK5; GB:AE004437; NID:g10581482; PIDN:MAG20215.1; GSPDB:G
 C:Genetics:
 A:Gene: VNG2064H

Query Match 35.0%; Score 7; DB 2; Length 571;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALQWR 11
 DB 28 LWALQWR 34

RESULT 5
 T46382
 hypothetical protein DKFZp434M0519.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
 C:Accession: T46382
 R:Ottenwälder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 Submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23031
 A:Accession: T46382
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-174 <AAA>
 A:Cross-references: UNIPROT:Q86T03; EMBL:AL137727
 A:Experimental source: adult testis; clone DKFZp434M0519
 C:Genetics:
 A:Note: DKFZp434M0519.1

Query Match 30.0%; Score 6; DB 2; Length 174;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLRR 15
|||||
Db 1 WRWLRR 6

RESULT 6

F82725
hypothetical protein XF1099 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82725
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: F82725
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-191 <SIM>
A:Cross-references: UNIPROT:Q9PEC9; GB:AE003945; GB:AE003849; NID:g9106036; PIDN:AAF8390
A:Experimental source: strain 9a5c
R:Simonsen, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H.; Netto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, E.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.B.; Laigle Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B.A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Teshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1099

Query Match 30.0%; Score 6; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
|||||
Db 94 SLWALG 99

RESULT 7

D83210
hypothetical protein PA3483 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83210
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83210
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-266 <STO>
A:Cross-references: UNIPROT:Q9HYC6; GB:AE004769; GB:AE004091; NID:g9949624; PIDN:AAG0687
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3483

Query Match 30.0%; Score 6; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLRR 15
|||||
Db 254 WRWLRR 259

RESULT 8

A83377
conserved hypothetical protein PA2154 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: A83377
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: A83377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-331 <STO>
A:Cross-references: UNIPROT:Q9I1W1; GB:AE004642; GB:AE004091; NID:g9948163; PIDN:AAG0554
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2154

Query Match 30.0%; Score 6; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLRR 15
|||||
Db 18 WRWLRR 23

RESULT 9

JN0530
thioesterase homolog - Streptomyces griseus
N:Alternate names: protein X
C:Species: Streptomyces griseus
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Jul-2004
C:Accession: JN0530
R:Criado, L.M.; Martin, J.F.; Gil, J.A.
Gene 126, 135-139, 1993
A:Title: The pab gene of Streptomyces griseus, encoding p-aminobenzoic acid synthase, is
A:Reference number: JN0530; MUID:93231527; PMID:8472954
A:Accession: JN0530
A:Molecule type: DNA
A:Residues: 1-361 <CRI>
A:Cross-references: UNIPROT:P33586; GB:M93058; NID:g153396; PIDN:AAA72110.1; PID:g388262
C:Superfamily: type II thioesterase with ferredoxin-like domain; oleoyl-lacyl-carrier-pr
F:129-339/Domain: oleoyl-lacyl-carrier-protein] hydrolase homology <ACPH>
F:201/Active site: Ser #status Predicted

Query Match 30.0%; Score 6; DB 2; Length 361;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RWLRRY 16
|||||
Db 115 RWLRRY 120

RESULT 10

S67856
GumG protein - Xanthomonas campestris
C:Species: Xanthomonas campestris
C>Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S67856
R:Pollock, T.J.
submitted to the EMBL Data Library, March 1995
A:Reference number: S67856

A:Accession: S67856
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-379 <POL>
 A:Cross-references: UNIPROT:Q56773; EMBL:U22511; NID:g1172090; PIDN:AAA86375.1; PID:g733
 C:Genetics:
 A:Gene: gumG
 C:Superfamily: O-acetyl transferase

Query Match 30.0%; Score 6; DB 2; Length 379;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GWRWLR 14
 |||||
 Db 279 GWRWLR 284

RESULT 11

F87644
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C:Accession: F87644
 R:Nieman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: F87644
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-411 <STO>
 A:Cross-references: UNIPROT:Q9A311; GB:AE005673; NID:g13424868; PIDN:AAK25154.1; GSPDB:G
 A:Gene: CC192

Query Match 30.0%; Score 6; DB 2; Length 411;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RWLRRY 16
 |||||
 Db 178 RWLRRY 183

RESULT 12

C83212
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: C83212
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
 J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: C83212
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-426 <STO>
 A:Cross-references: UNIPROT:Q9HYD1; GB:AE004768; GB:AE004091; NID:g9949611; PIDN:AAG0686
 A:Experimental source: strain PAO1
 C:Genetics:
 A:Gene: rhlB; PA3478
 C:Superfamily: Glycosyltransferase

Query Match 30.0%; Score 6; DB 2; Length 426;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||
 Db 106 SLWALG 111

RESULT 13

B53652
 C:Species: Pseudomonas aeruginosa
 C:Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 09-Jul-2004
 C:Accession: B53652
 R:Ochsner, U.A.; Fiechter, A.; Reiser, J.
 J. Biol. Chem. 269, 19787-19795, 1994
 A:Title: Isolation, characterization, and expression in Escherichia coli of the Pseudom
 A:Reference number: A53652; MUID:94327521; PMID:8051059
 A:Accession: B53652
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-426 <OCH>
 A:Cross-references: UNIPROT:Q51560; GB:L28170; NID:g452502; PIDN:AAA62129.1; PID:g452504
 C:Superfamily: glycosyltransferase

Query Match 30.0%; Score 6; DB 2; Length 426;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||
 Db 106 SLWALG 111

RESULT 14

S55273
 C:Species: Aspergillus niger
 C:Date: 14-Oct-1995 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
 C:Accession: S55273; S55267; S60356
 R:Schilling, B.; Lerch, K.
 Mol. Gen. Genet. 247, 430-438, 1995
 A:Title: Cloning, sequencing and heterologous expression of the monoamine oxidase gene f
 A:Reference number: S55267; MUID:95287865; PMID:7770050
 A:Accession: S55273
 A:Molecule type: DNA
 A:Residues: 1-495 <SCH>
 A:Cross-references: UNIPROT:P46882; EMBL:L38858; NID:g619754; PIDN:AAA98490.1; PID:g6197
 A:Accession: S55267
 A:Molecule type: protein
 A:Residues: 165-175 <SCW>
 R:Schilling, B.; Lerch, K.
 Biochim. Biophys. Acta 1243, 529-537, 1995
 A:Title: Amine oxidases from Aspergillus niger: identification of a novel flavin-depende
 A:Reference number: S60356; MUID:95244610; PMID:7727530
 A:Accession: S60356
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-52 <SCP>
 C:Genetics:
 A:Introns: 62/1; 89/3
 C:Keywords: FAD; flavoprotein; oxidoreductase; peroxisome
 F:41-69/Region: beta-alpha-beta FAD nucleotide-binding fold
 F:493-495/Region: peroxisome/glyoxysome location signal #status atypical

Query Match 30.0%; Score 6; DB 2; Length 495;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WALGWR 11
 |||||
 Db 459 WALGWR 464

RESULT 15

E86265
 protein P3F19.15 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: A86265
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.P.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86265
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-522 <STO>
 A:Cross-references: UNIPROT:Q9SAR6; GB:AE005172; NID:g4850396; PIDN:AAD31066.1; GSPDB:GN00002
 C:Genetics:
 A:Gene: F3F19.15
 A:Map position: 1

Query Match 30.0%; Score 6; DB 2; Length 522;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSLWAL 8
 |||||
 DB 319 WSLWAL 324

RESULT 16
 S46096
 probable membrane protein YBR220c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YBR1510
 C:Species: Saccharomyces cerevisiae
 C:Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
 C:Accession: S46096
 R:Dubois, E.; El Bakoury, M.; Glansdorff, N.; Messenguy, F.; Pierard, A.; Scherens, B.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45782
 A:Accession: S46096
 A:Molecule type: DNA
 A:Residues: 1-560 <DUB>
 A:Cross-references: UNIPROT:P38318; EMBL:Z36088; NID:g536609; PID:g536611; GSPDB:GN00002
 C:Genetics:
 A:Gene: MIPS:YBR220C
 A:Cross-references: SGD:S0000424
 A:Map position: 2R
 C:Keywords: transmembrane protein
 F:18-37/Domain: transmembrane #status predicted <TM01>
 F:98-106/Domain: transmembrane #status predicted <TM02>
 F:139-155/Domain: transmembrane #status predicted <TM03>
 F:177-193/Domain: transmembrane #status predicted <TM04>
 F:218-238/Domain: transmembrane #status predicted <TM05>
 F:330-347/Domain: transmembrane #status predicted <TM06>
 F:391-407/Domain: transmembrane #status predicted <TM07>
 F:421-437/Domain: transmembrane #status predicted <TM08>
 F:475-491/Domain: transmembrane #status predicted <TM09>
 F:521-544/Domain: transmembrane #status predicted <TM10>

Query Match 30.0%; Score 6; DB 2; Length 560;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALGW 10
 |||||
 DB 101 LWALGW 106

RESULT 17
 S76730
 hypothetical protein - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C:Accession: S76730
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis s.
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S76730
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-636 <KAN>
 A:Cross-references: UNIPROT:P74536; EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAAL864;
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 30.0%; Score 6; DB 2; Length 636;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LGWRWL 13
 |||||
 DB 52 LGWRWL 57

RESULT 18
 T02245
 hypothetical protein P1.11659.3 - human
 C:Species: Homo sapiens (man)
 C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 05-Nov-1999
 C:Accession: T02245
 R:Lamerdin, J.E.; McCreedy, P.M.; Skowronski, E.; Adamson, A.W.; Burkhardt-Schultz, K.; G Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Attix, C.; Andreise, T.; Trankheim submitted to the EMBL Data Library, March 1998
 A:Authors: Duarte, S.; Lucas, S.; Bruce, R.; Thomas, P.; Ouan, G.; Kronmiller, B.; Arella A:Description: Sequence analysis of a human P1 clone containing the XRCC9 DNA repair gene
 A:Reference number: Z14637
 A:Accession: T02245
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-880 <LAM>
 A:Cross-references: EMBL:AC004472; NID:g2984582; PIDN:AAC07985.1; PID:g2984587
 C:Genetics:
 A:Map position: 9
 A:Introns: 89/3; 152/1; 200/1; 241/2; 294/3; 354/3; 435/1; 733/1; 802/1; 871/1

Query Match 30.0%; Score 6; DB 2; Length 880;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRYG 17
 |||||
 DB 457 WLRRYG 462

RESULT 19
 F88559
 protein C48B4.4b [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 16-Aug-2004
 C:Accession: F88559
 R:anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.elegans
 A:Note: published errata appeared in Science 283, 35, 1999; Science 285, 2103, 1999; and A:Accession: F88559

```

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1758 <STO>
A:Cross-references: GB:chr_III; PIDN:CAA82384.1; PID:g3875025; GSPDB:GN00021; CESP:C48B4
C:Genetics:
A:Gene: C48B4.4b
A:Map position: 3
C:Superfamily: ATP-binding cassette homology

Query Match      30.0%; Score 6; DB 2; Length 1758;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RWLRRY 16
Db 16 RWLRRY 21

RESULT 20
S60124
transport protein homolog C48B4.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 13-Jan-1996 #sequence revision 12-Apr-1996 #text_change 16-Aug-2004
C:Accession: S60124; S40724; S40725
R:Kershaw, J.
submitted to the EMBL Data Library, November 1995
A:Reference number: S60124
A:Accession: S60124
A:Molecule type: DNA
A:Residues: 1-1767 <KER>
A:Cross-references: EMBL:229117; NID:g439247; PID:g1066912
C:Genetics:
A:Map position: III
A:Introns: 47/1; 112/3; 161/2; 220/2; 543/3; 574/3; 903/2; 1056/1; 1115/3; 1178/3; 1265/
C:Superfamily: ATP-binding cassette homology
C:Keywords: ATP; duplication; nucleotide binding; P-loop; transmembrane protein
F:628-818/Domain: ATP-binding cassette homology <ABC1>
F:645-652/Region: nucleotide-binding motif A (P-loop)
F:764-769/Region: nucleotide-binding motif B
F:1457-1642/Domain: ATP-binding cassette homology <ABC2>
F:1474-1481/Region: nucleotide-binding motif A (P-loop)
F:1586-1591/Region: nucleotide-binding motif B

Query Match      30.0%; Score 6; DB 2; Length 1767;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RWLRRY 16
Db 16 RWLRRY 21

RESULT 21
QXBP7L
hypothetical protein C-60 (nin region) - phage lambda
C:Species: phage lambda
C:Date: 13-Jun-1983 #sequence revision 13-Jun-1983 #text_change 09-Jul-2004
C:Accession: C43011; H43016; A04393
R:Danielis, D.
submitted to the Nucleic Acid Sequence Database, September 1982
A:Reference number: A94614
A:Accession: C43011
A:Molecule type: DNA
A:Residues: 1-60 <DAN>
A:Cross-references: UNIPROT:P03768
R:Sanger, F.; Coulson, A.R.; Hong, G.F.; Hill, D.F.; Petersen, G.B.
J. Mol. Biol. 162, 729-773, 1982
A:Title: Nucleotide sequence of bacteriophage lambda DNA.
A:Reference number: A92891; MUID:83189071; PMID:6221115
A:Accession: H43016
A:Molecule type: DNA
A:Residues: 1-60 <SAN>
A:Cross-references: GB:J02459; GB:M17233; GB:M24325; GB:V00636; GB:X00906; NID:g215104;

```

C:Genetics:

A:Map position: 86.78-87.15

A:Start codon: GTG

C:Superfamily: phage lambda hypothetical 7.4K protein (nin region)

Query Match

Best Local Similarity 25.0%; Score 5; DB 1; Length 60;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RWLRR 15

Db 50 RWLRR 54

RESULT 22

D83627

Na+/H+ antiporter nahs [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: D83827

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira-

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: D83827

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-65 <STO>

A:Cross-references: UNIPROT:Q9KCZ9; GB:AP001512; GB:BA000004; NID:gl0174030; PIDN:BA0051

A:Experimental source: strain C-125

C:Genetics:

A:Gene: nahs

Query Match

Best Local Similarity 25.0%; Score 5; DB 2; Length 65;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RWLRR 15

Db 33 RWLRR 37

RESULT 23

A72733

hypothetical protein APES012 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1993 #sequence_revision 20-Aug-1999 #text_change 03-Jul-2004

C:Accession: A72733

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah-

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: A72733

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-75 <KAW>

A:Cross-references: UNIPROT:Q9YF39; DDBJ:AP000059; NID:gs103911; PIDN:BAA79357.1; PID:d

A:Experimental source: strain K1

C:Genetics:

A:Gene: APES012

Query Match

Best Local Similarity 25.0%; Score 5; DB 2; Length 75;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LGWRW 12

Db 42 LGWRW 46

RESULT 24

S72700

probable ketoacyl synthase kas - Mycobacterium leprae
 N:Alternate names: Lep1170_Cl189 protein
 C:Species: Mycobacterium leprae
 C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C:Accession: S72700
 R:Smith, D.R.; Robison, K.
 submitted to the EMBL Data Library, November 1993
 A:Description: Mycobacterium leprae cosmid B1170.
 A:Reference number: S72693
 A:Accession: S72700
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-84 <SMI>
 A:Cross-references: UNIPROT:Q49621; EMBL:U00010; NID:g466780; PIDN:AAA17064.1; PID:g466780
 C:Genetics:
 A:Gene: kas
 A:Start codon: GTG

Query Match 25.0%; Score 5; DB 2; Length 84;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WALGW 10
 |||||
 Db 74 WALGW 78

RESULT 25
 I56009
 MHC HLA-A2-alpha-2 protein - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C:Accession: I56009
 R:Sire, J.; Chimini, G.; Boretto, J.; Toubert, A.; Kahn-Perles, B.; Layet, C.; Sodoyer, J. Immunol. 140, 2422-2430, 1988
 A>Title: Hybrid genes between HLA-A2 and HLA-A3 constructed by in vivo recombination all
 A:Reference number: I56009; MUID:88170828; PMID:2450922
 A:Accession: I56009
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-89 <RES>
 A:Cross-references: UNIPROT:Q30176; GB:M20139; NID:g188493; PIDN:AAA59838.1; PID:g553587
 C:Genetics:
 A:Gene: GDB:HLA-A
 A:Cross-references: GDB:119310; OMIM:142800
 A:Map position: 6p21.3-6p21.3
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 25.0%; Score 5; DB 2; Length 89;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
 |||||
 Db 74 WLRRY 78

RESULT 26
 I58989
 MHC H2-K transplantation antigen - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C:Accession: I58989
 R:Schulze, D.H.; Pease, L.R.; Geier, S.S.; Reyes, A.A.; Sarmiento, L.A.; Wallace, R.B.; Proc. Natl. Acad. Sci. U.S.A. 80, 2007-2011, 1983
 A>Title: comparison of the cloned h-2kbm1 variant gene with the h-2kb gene shows a clust
 A:Reference number: I58989; MUID:83169810; PMID:6300887
 A:Accession: I58989
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-91 <RES>
 A:Cross-references: UNIPROT:Q31190; GB:J00401; NID:g199540; PIDN:AAA39651.1; PID:g554232
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 25.0%; Score 5; DB 2; Length 91;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
 |||||
 Db 76 WLRRY 80

RESULT 27
 I57611
 MHC K-bm6 transplantation antigen - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
 C:Accession: I57611
 R:Gelliebert, J.; Zeff, R.A.; Schulze, D.H.; Pease, L.R.; Weiss, E.H.; Mellor, A.L.; Flave Mol. Cell. Biol. 6, 645-652, 1986
 A>Title: Interaction between K-b and Q4 gene sequences generates the K-bm6 mutation.
 A:Reference number: I57611; MUID:87064350; PMID:3023861
 A:Accession: I57611
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-91 <RES>
 A:Cross-references: UNIPROT:Q95555; GB:M12935; NID:g199595; PIDN:AAA39677.1; PID:g554238
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 25.0%; Score 5; DB 2; Length 91;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
 |||||
 Db 76 WLRRY 80

RESULT 28
 I59068
 MHC class I H2-K-b-alpha-2 cell surface glycoprotein - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C:Accession: I59068
 R:Gelliebert, J.; Zeff, R.A.; Melvold, R.W.; Nathenson, S.G. Proc. Natl. Acad. Sci. U.S.A. 83, 3371-3375, 1986
 A>Title: Mitotic recombination in germ cells generated two major histocompatibility comp
 A:Reference number: I59068; MUID:86205899; PMID:3458188
 A:Accession: I59068
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-91 <RES>
 A:Cross-references: UNIPROT:Q31189; GB:M13200; NID:g199535; PIDN:AAA39649.1; PID:g554230
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: glycoprotein

Query Match 25.0%; Score 5; DB 2; Length 91;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
 |||||
 Db 76 WLRRY 80

RESULT 29
 H30401
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - cutthroat trout mitochondrion (f
 C:Species: mitochondrion Salmo clarki (cutthroat trout)
 C>Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 09-Jul-2004
 C:Accession: H30401
 R:Thomas, W.K.; Beckenbach, A.T. J. Mol. Evol. 29, 233-245, 1989
 A>Title: Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanisms
 A:Reference number: A30396; MUID:89382658; PMID:2550657

A:Accession: H30401
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-92 <THO>
A:Cross-references: UNIPROT:P11630
C:Genetics:

A:Genome: mitochondrion
A:Genetic code: SGC1

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 25.0%; Score 5; DB 2; Length 92;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWAL 8

Db 45 SLWAL 49

RESULT 30

D30401

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - pink salmon mitochondrion (frag

C:Species: mitochondrion Oncorhynchus gorbuscha (pink salmon)

C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 09-Jul-2004

C:Accession: D30401

R:Thomas, W.K.; Beckenbach, A.T.

J. Mol. Evol. 29, 233-245, 1989

A:Title: Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanism

A:Reference number: A30396; MUID:89382658; PMID:2550657

A:Accession: D30401

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-92 <THO>

A:Cross-references: UNIPROT:P11630

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC1

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 25.0%; Score 5; DB 2; Length 92;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWAL 8

Db 45 SLWAL 49

RESULT 31

E30401

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - coho salmon mitochondrion (frag

C:Species: mitochondrion Oncorhynchus kisutch (coho salmon)

C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 09-Jul-2004

C:Accession: E30401

R:Thomas, W.K.; Beckenbach, A.T.

J. Mol. Evol. 29, 233-245, 1989

A:Title: Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanism

A:Reference number: A30396; MUID:89382658; PMID:2550657

A:Accession: E30401

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-92 <THO>

A:Cross-references: UNIPROT:P11630

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC1

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 25.0%; Score 5; DB 2; Length 92;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWAL 8

Db 45 SLWAL 49

RESULT 32

F30401

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - sockeye salmon mitochondrion (f

C:Species: mitochondrion Oncorhynchus nerka (sockeye salmon)

C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 09-Jul-2004

C:Accession: F30401

R:Thomas, W.K.; Beckenbach, A.T.

J. Mol. Evol. 29, 233-245, 1989

A:Title: Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanism

A:Reference number: A30396; MUID:89382658; PMID:2550657

A:Accession: F30401

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-92 <THO>

A:Cross-references: UNIPROT:P11630

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC1

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 25.0%; Score 5; DB 2; Length 92;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWAL 8

Db 45 SLWAL 49

RESULT 33

G30401

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - chinook salmon mitochondrion (f

C:Species: mitochondrion Oncorhynchus tshawytscha (chinook salmon)

C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 09-Jul-2004

C:Accession: G30401

R:Thomas, W.K.; Beckenbach, A.T.

J. Mol. Evol. 29, 233-245, 1989

A:Title: Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanism

A:Reference number: A30396; MUID:89382658; PMID:2550657

A:Accession: G30401

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-92 <THO>

A:Cross-references: UNIPROT:P11630

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC1

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 25.0%; Score 5; DB 2; Length 92;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWAL 8

Db 45 SLWAL 49

RESULT 34

S45490

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - European seabass mitochondrion

C:Species: mitochondrion Dicentrarchus labrax (European seabass)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C:Accession: S45490

R;Venanzetti, F.; Cecconi, F.; Giorgi, M.; Cesaroni, D.; Sbordoni, V.; Mariottini, P.
 Curr. Genet. 26, 139-145, 1994
 A;Title: Cloning and characterization of the European seabass, Dicentrarchus labrax, mit
 A;Reference number: S45489; MUID:95094310; PMID:8001168
 A;Accession: S45490
 A;Molecule type: DNA
 A;Residues: 1-97 <VEN>
 A;Cross-references: UNIPROT:Q36364; EMBL:X74148; NID:g521080; PIDN:CAA52245.1; PID:g5210
 C;Genetics:
 A;Genome: mitochondrion
 A;Genetic code: SGC1
 C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L
 C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
 Query Match 25.0%; Score 5; DB 2; Length 97;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWAL 8
 Db 44 SLWAL 48
 RESULT 35
 T09865
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - rainbow trout mitochondrion
 C;Species: mitochondrion Oncorhynchus mykiss (rainbow trout)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C;Accession: T09865; C30401; S04689
 R;Zardoya, R.; Garrido-Perterra, A.; Bautista, J.M.
 J. Mol. Evol. 41, 942-951, 1995
 A;Title: The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow
 A;Reference number: Z16890; MUID:96139027; PMID:8587139
 A;Accession: T09865
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-98 <ZAR>
 A;Cross-references: UNIPROT:P11630; EMBL:L29771; NID:g1246865; PIDN:AAB03355.1; PID:g463
 A;Experimental source: liver
 R;Thomas, W.K.; Beckenbach, A.T.
 J. Mol. Evol. 29, 233-245, 1989
 A;Title: Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanism
 A;Reference number: A30396; MUID:89382658; PMID:2550657
 A;Accession: C30401
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-92 <THO>
 R;Davidson, W.S.; Bartlett, S.E.; Birt, T.P.; Green, J.M.
 Curr. Genet. 14, 483-486, 1988
 A;Title: Cloning and sequence analysis of an XbaI fragment of rainbow trout mitochondria
 A;Reference number: S04688; MUID:89136066; PMID:3265651
 A;Accession: S04689
 A;Molecule type: DNA
 A;Residues: 1-2, 'ALMHQYTS', 8-25, 'FS', 28, 'P', 30-35 <DAV>
 A;Cross-references: EMBL:X14013; NID:g13622; PIDN:CAA32179.1; PID:g13624
 C;Genetics:
 A;Gene: NADH-4L
 A;Genome: mitochondrion
 A;Genetic code: SGC1
 C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L
 C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
 Query Match 25.0%; Score 5; DB 2; Length 98;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWAL 8
 Db 45 SLWAL 49
 RESULT 36
 T09955

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - Atlantic salmon mitochondrion
 C;Species: mitochondrion Salmo salar (Atlantic salmon)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C;Accession: T09955
 R;Hurst, C.D.; Bartlett, S.E.; Bruce, I.J.; Davidson, W.S.
 submitted to the EMBL Data Library, October 1998
 A;Description: The complete nucleotide sequence of the mitochondrial DNA of the Atlantic
 A;Reference number: Z16904
 A;Accession: T09955
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-98 <HUR>
 A;Cross-references: UNIPROT:Q922M5; EMBL:U12143; NID:g3775976; PID:g3776543
 C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L
 C;Keywords: electron transfer; membrane-associated complex; mitochondrion; NAD; oxidative
 Query Match 25.0%; Score 5; DB 2; Length 98;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWAL 8
 Db 45 SLWAL 49
 RESULT 37
 JQ0906
 bombyxin A-3 homolog - ailanthus silkmoth
 C;Species: Samia cynthia (ailanthus silkmoth)
 C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
 C;Accession: JQ0906
 R;Iwami, M.
 submitted to JIPID, March 1991
 A;Reference number: JQ0902
 A;Accession: JQ0906
 A;Molecule type: DNA
 A;Residues: 1-98 <IWA>
 A;Cross-references: UNIPROT:P33720
 A;Note: the authors translated the codon ATA for residue 32 as Thr
 A;Note: the source is designated as Samia cynthia ricini
 C;Genetics:
 A;Gene: sbxA3
 C;Superfamily: insulin
 C;Keywords: hormone
 Query Match 25.0%; Score 5; DB 2; Length 98;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 GWRWL 13
 Db 58 GWRWL 62
 RESULT 38
 JQ0902
 bombyxin A-1 homolog - ailanthus silkmoth
 C;Species: Samia cynthia (ailanthus silkmoth)
 C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
 C;Accession: JQ0902
 R;Iwami, M.
 submitted to JIPID, March 1991
 A;Reference number: JQ0902
 A;Accession: JQ0902
 A;Molecule type: DNA
 A;Residues: 1-99 <IWA>
 A;Cross-references: UNIPROT:P33718
 A;Note: the source is designated as Samia cynthia ricini

QY 11 RWLRR 15

Arch. Biochem. Biophys.

A;Title: The novel ion pump rhodopsins from *Halocarcula* form a family independent from bo

A;Reference number: S51206; MUID:95070153; PMID:7979388

A;Accession: S51207

A;Molecule type: DNA

A;Residues: 1-118 <TAT>

A;Experimental source: strain arg-1

C;Superfamily: bacteriorhodopsin

C;Keywords: chromoprotein; photoreceptor; retinal; transmembrane protein

Query Match 25.0%; Score 5; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9

|||||

Db 90 LWALG 94

RESULT 44

D97046 hypothetical protein CAC1187 [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004

C;Accession: D97046

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J; Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: D97046

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-118 <KUP>

A;Cross-references: UNIPROT:Q97J76; GB:AE001437; PIDN:AAK79159.1; PID:g15024109; GSPDB:G

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC1187

Query Match

Best Local Similarity 25.0%; Score 5; DB 2; Length 118;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRYG 17

|||||

Db 13 LRRYG 17

RESULT 45

B71020

hypothetical protein PH1455 - Pyrococcus horikoshii

C;Species: Pyrococcus horikoshii

C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004

C;Accession: B71020

R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: B71020

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-120 <RAW>

A;Cross-references: UNIPROT:O59124; GB:AP000006; NID:g3236133; PIDN:BAA30562.1; PID:g323

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:

A;Gene: PH1455

Query Match

Best Local Similarity 25.0%; Score 5; DB 2; Length 120;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9

|||||

Db 56 LWALG 60

RESULT 46

T05357

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) 14K chain - Arabidopsis thaliana

N;Alternate names: protein F8B4.170

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 27-Oct-2003

C;Accession: T05357

R;Bevan, M.; Terryn, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De P

ewes, H.W.; Mayer, K.F.X.; Schueller, C.

submitted to the Protein Sequence Database, February 1999

A;Reference number: Z15409

A;Accession: T05357

A;Molecule type: DNA

A;Residues: 1-122 <BEV>

A;Cross-references: EMBL:AL034567

A;Experimental source: cultivar Columbia; BAC clone F8B4

C;Genetics:

A;Map position: 4

A;Genome: nuclear

A;Introns: 32/1; 84/3; 100/3

A;Note: F8B4.170

C;Superfamily: ubiquinol-cytochrome-c reductase 14 kDa protein

C;Keywords: electron transfer; membrane protein; membrane-associated complex; mitochondri

Query Match

Best Local Similarity 25.0%; Score 5; DB 2; Length 122;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRYG 17

|||||

Db 28 LRRYG 32

RESULT 47

T03359

gene e8 protein - Lactococcus phage b11170

C;Species: Lactococcus phage b11170

C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004

C;Accession: T03359

R;Crutz-Le Coq, A.M.; Cesselin, B.; Commissaire, J.; Anba, J.; Kyriakidis, S.; Chopin, M.

submitted to the EMBL Data Library, June 1997

A;Description: Sequence and organization of the lactococcal isometric b11170 phage genome

A;Reference number: Z14903

A;Accession: T03359

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-125 <CRU>

A;Cross-references: UNIPROT:O80152; EMBL:AF009630; NID:g3282260; PIDN:AAC27230.1; PID:g32

C;Genetics:

A;Gene: e8

C;Superfamily: Lactococcus phage b11170 gene e8 protein

Query Match

Best Local Similarity 25.0%; Score 5; DB 2; Length 125;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 RYGNW 19

|||||

Db 27 RYGNW 31

RESULT 48

I63123

cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Geomys brevicaeps) mitochondri

C;Species: mitochondrion Geomys brevicaeps (pocket gopher)

C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004

C;Accession: I63123

R;Haefner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler,

Science 265, 1087-1090, 1994

A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.

A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I63123
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q34550; GB:L32683; NID:G548223; PIDN:AAA66279.1; PID:G548224
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metalloprote

Query Match 25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
DB 98 LWALG 102

RESULT 49
I48089
cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Cratogeomys merriami) mitochondrion
C;Species: mitochondrion Cratogeomys merriami (pocket gopher)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I48089
R;Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler
Science 265, 1087-1090, 1994
A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I48089
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q34162; GB:L32688; NID:G548110; PIDN:AAA66244.1; PID:G548111
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metalloprote

Query Match 25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
DB 98 LWALG 102

RESULT 50
I63124
cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Cratogeomys bulleri) mitochondrion
C;Species: mitochondrion Cratogeomys bulleri (pocket gopher)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I63124
R;Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler
Science 265, 1087-1090, 1994
A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I63124
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q35388; GB:L32684; NID:G548225; PIDN:AAA66280.1; PID:G548226
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metalloprote

Query Match 25.0%; Score 5; DB 2; Length 126;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
DB 98 LWALG 102

RESULT 51
I63121
cytochrome-c oxidase (EC 1.9.3.1) chain I - western pocket gopher (Thomomys bottae) mitochondrion
C;Species: mitochondrion Thomomys bottae (western pocket gopher)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I63121
R;Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler
Science 265, 1087-1090, 1994
A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I63121
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q36004; GB:L32682; NID:G548217; PIDN:AAA66277.1; PID:G548218
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metalloprote

Query Match 25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
DB 98 LWALG 102

RESULT 52
I63125
cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Geomys personatus) mitochondrion
C;Species: mitochondrion Geomys personatus (pocket gopher)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I63125
R;Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler
Science 265, 1087-1090, 1994
A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I63125
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q34642; GB:L32689; NID:G548227; PIDN:AAA66281.1; PID:G548228
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metalloprote

Query Match 25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
DB 98 LWALG 102

RESULT 53
I63122
cytochrome-c oxidase (EC 1.9.3.1) chain I - western pocket gopher (Thomomys talpoides) mitochondrion
C;Species: mitochondrion Thomomys talpoides (western pocket gopher)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004

C;Accession: I63122
R;Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler, Science 265, 1087-1090, 1994
A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I63122
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q36172; UNIPROT:Q8M605; UNIPROT:Q8M601; GB:L32690; NID:G548204
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metallopro

Query Match 25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
Db 98 LWALG 102

RESULT 54
I63116
Cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Orthogeomys cavator) mitochondrion
C;Species: mitochondrion Orthogeomys cavator (pocket gopher)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I63116
R;Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler, Science 265, 1087-1090, 1994
A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I63116
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q35240; GB:L32686; NID:G548197; PIDN:AAA66272.1; PID:G548198
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metallopro

Query Match 25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
Db 98 LWALG 102

RESULT 55
I63117
Cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Orthogeomys cherriei) mitochondrion
C;Species: mitochondrion Orthogeomys cherriei (pocket gopher)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I63117
R;Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler, Science 265, 1087-1090, 1994
A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I63117
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q35241; GB:L32687; NID:G548199; PIDN:AAA66273.1; PID:G548200
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI

C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metallopro

Query Match 25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
Db 98 LWALG 102

RESULT 56
I63120
Cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Orthogeomys heterodus) mitochondrion
C;Species: mitochondrion Orthogeomys heterodus (pocket gopher)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I63120
R;Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler, Science 265, 1087-1090, 1994
A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I63120
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q35257; GB:L32696; NID:G548205; PIDN:AAA66276.1; PID:G548206
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metallopro

Query Match 25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
Db 98 LWALG 102

RESULT 57
I63119
Cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Orthogeomys hispidus) mitochondrion
C;Species: mitochondrion Orthogeomys hispidus (pocket gopher)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I63119
R;Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler, Science 265, 1087-1090, 1994
A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I63119
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q35256; GB:L32692; NID:G548203; PIDN:AAA66275.1; PID:G548204
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metallopro

Query Match 25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
Db 98 LWALG 102

RESULT 58
I63118
Cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Orthogeomys hispidus) mitochondrion
C;Species: mitochondrion Orthogeomys hispidus (pocket gopher)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I63118
R;Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler, Science 265, 1087-1090, 1994
A;Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
A;Reference number: I48088; MUID:94345374; PMID:8066445
A;Accession: I63118
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-126 <RES>
A;Cross-references: UNIPROT:Q35255; GB:L32693; NID:G548202; PIDN:AAA66274.1; PID:G548203
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: copper; electron transfer; magnesium; membrane-associated complex; metallopro

Query Match 25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
Db 98 LWALG 102

I63118
 cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Orthogeomys underwoodi) mitochondrion
 C:Species: mitochondrion Orthogeomys underwoodi (pocket gopher)
 C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
 C:Accession: I63118
 R:Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler Science 265, 1087-1090, 1994
 A:Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
 A:Reference number: I48088; PMID:94345374; PMID:8066445
 A:Accession: I63118
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-126 <RES>
 A:Cross-references: UNIPROT:Q35337; GB:L32691; NID:G548201; PIDN:AAA66274.1; PID:G548202
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C:Keywords: copper; electron transfer; magnesium; membrane-associated complex; metalloprote

Query Match 25.0%; Score 5; DB 2; Length 126;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LWALG 9
 |||||
 Db 98 LWALG 102

RESULT 59
 I63128
 cytochrome-c oxidase (EC 1.9.3.1) chain I - Zygoeomys trichopus mitochondrion (fragment)
 C:Species: mitochondrion Zygoeomys trichopus
 C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
 C:Accession: I63128
 R:Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler Science 265, 1087-1090, 1994
 A:Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
 A:Reference number: I48088; PMID:94345374; PMID:8066445
 A:Accession: I63128
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-126 <RES>
 A:Cross-references: UNIPROT:Q36287; GB:L32695; NID:G548333; PIDN:AAA66324.1; PID:G548334
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C:Keywords: copper; electron transfer; magnesium; membrane-associated complex; metallopro

Query Match 25.0%; Score 5; DB 2; Length 126;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LWALG 9
 |||||
 Db 98 LWALG 102

RESULT 60
 I63126
 cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Geomys bursarius) mitochondrion
 C:Species: mitochondrion Geomys bursarius (pocket gopher)
 C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
 C:Accession: I63126
 R:Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler Science 265, 1087-1090, 1994
 A:Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
 A:Reference number: I48088; PMID:94345374; PMID:8066445
 A:Accession: I63126
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-126 <RES>

A:Cross-references: UNIPROT:Q34548; GB:L32693; NID:G548229; PIDN:AAA66282.1; PID:G548230
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C:Keywords: copper; electron transfer; magnesium; membrane-associated complex; metallopro

Query Match 25.0%; Score 5; DB 2; Length 126;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LWALG 9
 |||||
 Db 98 LWALG 102

RESULT 61
 I63127
 cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Geomys bursarius) mitochondrion
 C:Species: mitochondrion Geomys bursarius (pocket gopher)
 C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
 C:Accession: I63127
 R:Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler Science 265, 1087-1090, 1994
 A:Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
 A:Reference number: I48088; PMID:94345374; PMID:8066445
 A:Accession: I63127
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-126 <RES>
 A:Cross-references: UNIPROT:Q34549; GB:L32694; NID:G548231; PIDN:AAA66283.1; PID:G548232
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C:Keywords: copper; electron transfer; magnesium; membrane-associated complex; metallopro

Query Match 25.0%; Score 5; DB 2; Length 126;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LWALG 9
 |||||
 Db 98 LWALG 102

RESULT 62
 I48088
 cytochrome-c oxidase (EC 1.9.3.1) chain I - pocket gopher (Cratogeomys castaneops) mitochondrion
 C:Species: mitochondrion Cratogeomys castaneops (pocket gopher)
 C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
 C:Accession: I48088
 R:Hafner, M.S.; Sudman, P.D.; Villablanca, F.X.; Spradling, T.A.; Demastes, J.W.; Nadler Science 265, 1087-1090, 1994
 A:Title: Disparate rates of molecular evolution in cospeciating hosts and parasites.
 A:Reference number: I48088; PMID:94345374; PMID:8066445
 A:Accession: I48088
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-126 <RES>
 A:Cross-references: UNIPROT:Q34074; GB:L32685; NID:G548106; PIDN:AAA66243.1; PID:G548109
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C:Keywords: copper; electron transfer; magnesium; membrane-associated complex; metallopro

Query Match 25.0%; Score 5; DB 2; Length 126;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LWALG 9
 |||||

```
Db          98 LWALG 102

RESULT 63
S77183
hypothetical protein sll1698 - Synecocystis sp. (strain PCC 6803)
C/Species: Synecocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S77183
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
S.
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S77183
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-126 <KAN>
A/Cross-references: UNIPROT:P73694; EMBL:D90908; GB:AB001339; NID:91652725; PIDN:BAAL1774
A/Note: the nucleotide sequence was submitted to the EMBL data Library, June 1996
C/Superfamily: Synecocystis hypothetical protein sll1698

Query Match          25.0%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 LWALG 9
          |||||
Db          22 LWALG 26

RESULT 64
F45345
trans-regulatory splicing-like protein - caprine arthritis-encephalitis virus (strain CC
N/Alternate names: rev-like protein
C/Species: caprine arthritis-encephalitis virus, CAEV
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: F45345
R;Saltarelli, M.; Querat, G.; Konings, D.A.M.; Vigne, R.; Clements, J.E.
Virology 179, 347-364, 1990
A/Title: Nucleotide sequence and transcriptional analysis of molecular clones of CAEV wh
A/Reference number: A45345; MUID:91021037; PMID:2171210
A/Accession: F45345
A/Molecule type: mRNA
A/Residues: 1-133 <SAL>
A/Cross-references: UNIPROT:P33460; GB:M33677
C/Genetics:
A/Gene: rev
A/Introns: 38/1
C/Superfamily: caprine arthritis-encephalitis virus rev-like protein
C/Keywords: splicing protein; transcription regulation

Query Match          25.0%; Score 5; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          10 WRWLR 14
          |||||
Db          68 WRWLR 72

RESULT 65
VKLJCE
trans-regulatory splicing-like protein - caprine arthritis-encephalitis virus
N/Alternate names: rev-like protein
C/Species: caprine arthritis-encephalitis virus, CAEV
C/Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C/Accession: A40479
R;Kalinski, H.; Yaniv, A.; Mashiah, P.; Miki, T.; Tronick, S.R.; Gazit, A.
Virology 183, 786-792, 1991
A/Title: rev-like transcripts of caprine arthritis encephalitis virus.

A/Reference number: A40479; MUID:91306466; PMID:1649509
A/Accession: A40479
A/Molecule type: mRNA
A/Residues: 1-133 <KAL>
A/Cross-references: UNIPROT:P31628; GB:M63105; NID:9323291; PIDN:AAA42893.1; PID:g323292
C/Genetics:
A/Gene: rev
C/Superfamily: caprine arthritis-encephalitis virus rev-like protein
C/Keywords: splicing protein; transcription regulation

Query Match          25.0%; Score 5; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          10 WRWLR 14
          |||||
Db          68 WRWLR 72

RESULT 66
S57492
cytochrome-c oxidase (EC 1.9.3.1) chain I - Carabus violaceus mitochondrion (fragment)
N/Alternate names: cytochrome oxidase I
C/Species: mitochondrion Carabus violaceus
C/Date: 19-Mar-1997 #sequence_revision 24-Oct-1997 #text_change 09-Jul-2004
C/Accession: S57492
R;Howland, D.E.; Hewitt, G.M.
submitted to the EMBL Data Library, May 1995
A/Description: Phylogeny of the Coleoptera based on mitochondrial cytochrome oxidase I se
A/Reference number: S57491
A/Accession: S57492
A/Molecule type: DNA
A/Residues: 1-133 <HOW>
A/Cross-references: UNIPROT:Q36356; EMBL:X88939; NID:9872231; PIDN:CAA61377.1; PID:g12892
C/Genetics:
A/Gene: COI
A/Genome: mitochondrion
A/Genetic code: SGC4
C/Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C/Keywords: chromoprotein; copper; electron transfer; heme; iron; magnesium; membrane-as
F;90/Binding site: magnesium (His) (shared with chain II) #status predicted
F;98/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match          25.0%; Score 5; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 LWALG 9
          |||||
Db          61 LWALG 65

RESULT 67
S57491
cytochrome-c oxidase (EC 1.9.3.1) chain I - Lucanus cervus mitochondrion (fragment)
N/Alternate names: cytochrome oxidase I
C/Species: mitochondrion Lucanus cervus
C/Date: 19-Mar-1997 #sequence_revision 24-Oct-1997 #text_change 09-Jul-2004
C/Accession: S57491
R;Howland, D.E.; Hewitt, G.M.
submitted to the EMBL Data Library, May 1995
A/Description: Phylogeny of the Coleoptera based on mitochondrial cytochrome oxidase I se
A/Reference number: S57491
A/Accession: S57491
A/Molecule type: DNA
A/Residues: 1-133 <HOW>
A/Cross-references: UNIPROT:Q36418; EMBL:X88954; NID:9872245; PIDN:CAA61391.1; PID:g12892
C/Genetics:
A/Gene: COI
A/Genome: mitochondrion
A/Genetic code: SGC4
C/Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C/Keywords: chromoprotein; copper; electron transfer; heme; iron; magnesium; membrane-as
```

F:90/Binding site: magnesium (His) (shared with chain II) #status predicted
F:98/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 25.0%; Score 5; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
Db 61 LWALG 65

RESULT 68
A10999
conserved hypothetical protein yrfA [imported] - Salmonella enterica subsp. enterica ser.
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: A10999
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerston, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: A10999
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-133 <PAR>
A:Cross-references: GB:AJ513382; PIDN:CAD08125.1; PID:gl6505104; GSPDB:GN00176
C:Genetics:
A:Gene: yrfA
C:Superfamily: Escherichia coli hypothetical 16.9K protein b3392

Query Match 25.0%; Score 5; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLR 14
|||||
Db 97 WRWLR 101

RESULT 69
D5511
bphF 3'-region hypothetical protein - Pseudomonas sp. (strain KKS012)
C:Species: Pseudomonas sp.
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 20-Jun-2000
C:Accession: D5511
R:Kikuchi, Y.; Yasukochi, Y.; Nagata, Y.; Fukuda, M.; Takagi, M.
J. Bacteriol. 176, 4269-4276, 1994
A:Title: Nucleotide sequence and functional analysis of the meta-cleavage pathway involv
A:Reference number: A55511; MUID:94292455; PMID:8021212
A:Accession: D5511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-136 <KIK>
A:Cross-references: GB:D16407; NID:g303760; PIDN:BAA03894.1; PID:g425215

Query Match 25.0%; Score 5; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
Db 8 LWALG 12

RESULT 70
I80172
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)

F:90/Binding site: magnesium (His) (shared with chain II) #status predicted
F:98/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 25.0%; Score 5; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
Db 61 LWALG 65

RESULT 68
A10999
conserved hypothetical protein yrfA [imported] - Salmonella enterica subsp. enterica ser.
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: A10999
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerston, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: A10999
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-133 <PAR>
A:Cross-references: GB:AJ513382; PIDN:CAD08125.1; PID:gl6505104; GSPDB:GN00176
C:Genetics:
A:Gene: yrfA
C:Superfamily: Escherichia coli hypothetical 16.9K protein b3392

Query Match 25.0%; Score 5; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLR 14
|||||
Db 97 WRWLR 101

RESULT 69
D5511
bphF 3'-region hypothetical protein - Pseudomonas sp. (strain KKS012)
C:Species: Pseudomonas sp.
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 20-Jun-2000
C:Accession: D5511
R:Kikuchi, Y.; Yasukochi, Y.; Nagata, Y.; Fukuda, M.; Takagi, M.
J. Bacteriol. 176, 4269-4276, 1994
A:Title: Nucleotide sequence and functional analysis of the meta-cleavage pathway involv
A:Reference number: A55511; MUID:94292455; PMID:8021212
A:Accession: D5511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-136 <KIK>
A:Cross-references: GB:D16407; NID:g303760; PIDN:BAA03894.1; PID:g425215

Query Match 25.0%; Score 5; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LWALG 9
|||||
Db 8 LWALG 12

RESULT 70
I80172
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)

C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C:Accession: I80172
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544; PMID:8016085
A:Accession: I80172
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: UNIPROT:Q95531; EMBL:U05583; NID:g454783; PIDN:AAA50186.1; PID:g4547
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 25.0%; Score 5; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
|||||
Db 132 WLRRY 136

RESULT 71
I80175
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C:Accession: I80175
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544; PMID:8016085
A:Accession: I80175
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-137 <RES>
A:Cross-references: UNIPROT:Q95534; EMBL:U05586; NID:g454789; PIDN:AAA50189.1; PID:g4547
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 25.0%; Score 5; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
|||||
Db 132 WLRRY 136

RESULT 72
I80173
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C:Accession: I80173
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544; PMID:8016085
A:Accession: I80173
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: UNIPROT:Q95532; EMBL:U05584; NID:g454785; PIDN:AAA50187.1; PID:g4547
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 25.0%; Score 5; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRY 16
|||||
Db 132 WLRRY 136

```
RESULT 73
I80176
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C>Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C:Accession: I80176
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544; PMID:8016085
A:Accession: I80176
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-137 <RES>
A:Cross-references: UNIPROT:Q95535; EMBL:U05587; NID:9454791; PIDN:AAA50190.1; PID:94547
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match      25.0%; Score 5; DB 2; Length 137;
Best Local Similarity 100.0%; Pred.No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      12 WLRRY 16
      |||||
Db      132 WLRRY 136

RESULT 74
I80174
class I histocompatibility antigen - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C>Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C:Accession: I80174
R:McAdam, S.N.; Boyson, J.E.; Liu, X.; Garber, T.L.; Hughes, A.L.; Bontrop, R.E.; Watkin
Proc. Natl. Acad. Sci. U.S.A. 91, 5893-5897, 1994
A:Title: A uniquely high level of recombination at the HLA-B locus.
A:Reference number: I59308; MUID:94286544; PMID:8016085
A:Accession: I80174
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: UNIPROT:Q95533; EMBL:U05585; NID:9454787; PIDN:AAA50188.1; PID:94547
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match      25.0%; Score 5; DB 2; Length 137;
Best Local Similarity 100.0%; Pred.No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      12 WLRRY 16
      |||||
Db      132 WLRRY 136

RESULT 75
I38875
MHC class I antigen - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004
C:Accession: I38875
R:Garber, T.L.; Butler, L.M.; Trachtenberg, E.A.; Erlich, H.A.; Rickards, O.; De Stefano
Immunogenetics 42, 19-27, 1995
A:Title: HLA-B alleles of the Cayapa of Ecuador: new B39 and B15 alleles.
A:Reference number: I38860; MUID:95317819; PMID:7797264
A:Accession: I38875
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: UNIPROT:Q29664; EMBL:U15639; NID:G930332; PIDN:AAA74046.1; PID:G9303
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match      25.0%; Score 5; DB 2; Length 137;
Best Local Similarity 100.0%; Pred.No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 12 WLRRY 16
 |||||
Db 132 WLRRY 136

Search completed: October 26, 2004, 07:23:31
Job time : 19.5 secs

This Page Blank (uspto)


```

RESULT 1
Q910H7
ID Q910H7 PRELIMINARY; PRT; 297 AA.
AC Q910H7;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PA2661;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey C.K., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004695; AAG06049.1; -.
DR FIR; CB3311; CB3311.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 297 AA; 33197 MW; 0DF94B92126D72D5 CRC64;

Query Match 40.0%; Score 8; DB 2; Length 297;
Best Local Similarity 100.0%; Pred.No. 0.67; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0;

QY 8 LGWRWLR 15
Db 72 LGWRWLR 79

RESULT 2
Q8UD12
ID Q8UD12 PRELIMINARY; PRT; 504 AA.
AC Q8UD12;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-WAR-2004 (TReMBLrel. 26, Last annotation update)
DE ABC transporter, membrane spanning protein.
GN OrderedLocusNames=Atu2143;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Serubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Bisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323 (2001).
DR EMBL; AE009164; AAL43133.1; -.
DR FIR; A97617; A97617.

```

```

DR FIR; AG2839; AG2839.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002293; AA/rel_permeasel.
DR InterPro; IPR000005; HTHArAc.
DR InterPro; IPR004841; Permease region.
DR Pfam; PF00324; AA_permease; 1_
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 504 AA; 54004 MW; 2028A16BH07D2461 CRC64;

Query Match 40.0%; Score 8; DB 2; Length 504;
Best Local Similarity 100.0%; Pred.No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VWSLWALG 9
Db 29 VWSLWALG 36

RESULT 3
Q7CX34
ID Q7CX34 PRELIMINARY; PRT; 518 AA.
AC Q7CX34;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE AGR_C_3885p.
GN OrderedLocusNames=AGR_C_3885;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cereon;
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurrello B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328 (2001).
DR EMBL; AE008130; AAK87890.1; -.
DR InterPro; IPR002293; AA/rel_permeasel.
DR InterPro; IPR000005; HTHArAc.
DR Pfam; PF00324; AA_permease; 1_
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
KW Transmembrane; Transport.
SQ SEQUENCE 518 AA; 55586 MW; 032E457438EB8555 CRC64;

Query Match 40.0%; Score 8; DB 2; Length 518;
Best Local Similarity 100.0%; Pred.No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VWSLWALG 9
Db 43 VWSLWALG 50

RESULT 4
Q7NUL1
ID Q7NUL1 PRELIMINARY; PRT; 351 AA.
AC Q7NUL1;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)

```

DE Protable ABC transporter, permease protein.
GN OrderedLocusNames=CV2686;
OS Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12472 / DSM 30191;
RX MEDLINE=22882880; PubMed=14500782;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.P.,
RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
RA Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Burity H.A.,
RA Camargo A.A., Cardoso D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.M.B., Cascado J.C.M., Cavada B.S., Chueire L.M.O.,
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
RA Grattapaglia D., Grissard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leoni L.C.F., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA Madeira H.M.F., Manlio G.P., Maranhao A.Q., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros S.R.B., Weissner R.V., Moreira M.A.M.,
RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seanez H.N.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steindl M., Teixeira S.R., Urmenyi T.,
RA Vettore A., Wassem R., Zaha A., Simpson A.J.G.;
RA "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
RC -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
CC EMBL: AE016919; AAQ60356.1; -;
DR GO: GO:0016020; C-membrane; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000515; BPD transp.
DR Pfam: PF00528; BPD transp 1; 1.
DR PROSITE: PS00928; ABC TM1; 1.
DR Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 351 AA; 38472 MW; CB208BE6ACAC689 CRC64;
Query Match 35.0%; Score 7; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 SLWALGW 10
Db 19 SLWALGW 25
RESULT 5
Q9HNK5 PRELIMINARY; PRT; 571 AA.
AC Q9HNK5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Vng2064h.
GN OrderedLocusNames=VNG2064H;
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).

OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leitnauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL: AB005097; AAG20215.1; -;
DR FIR; C84356; C84356.
KW Complete proteome.
SQ SEQUENCE 571 AA; 61059 MW; ED66FC0A7D35C176 CRC64;
Query Match 35.0%; Score 7; DB 2; Length 571;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 LWALGWR 11
Db 28 LWALGWR 34
RESULT 6
Q88G85 PRELIMINARY; PRT; 74 AA.
AC Q88G85;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Hypothetical protein.
GN OrderedLocusNames=PP3840;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazed A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Medler H., Lauber J., Stjepandic D., Hohlseil J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL: AE016788; AAN69434.1; -;
DR TIGR; PP3840; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 74 AA; 8550 MW; 5F7FEBF63468BD4 CRC64;
Query Match 30.0%; Score 6; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 WLRRYG 17
Db 69 WLRRYG 74
RESULT 7

```
Q6YXZ8
ID Q6YXZ8 PRELIMINARY; PRT; 100 AA.
AC Q6YXZ8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSUNBB0056122.58 (Hypothetical protein
DE OSUNBB0046012.18).
GN Name=OSJNB0056122.58; Synonyms=OSJNB0046012.18;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP005644; BADI7502.1; -.
DR EMBL; AP005643; BADI7451.1; -.
KW Hypothetical protein.
SQ SEQUENCE 100 AA; 10960 MW; CC21DOB5AFBD0E4D CRC64;

Query Match 30.0%; Score 6; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLRR 15
DB 19 WRWLRR 24

RESULT 8
Q8G7W5 PRELIMINARY; PRT; 100 AA.
ID Q8G7W5;
AC Q8G7W5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=BL0120;
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787;
RA Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AF014627; AAN23985.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003425; Unk_YGGT.
DR Pfam; PF02325; YGGT; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 100 AA; 11654 MW; 7B219E7C9B016BF9 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RWLRRY 16
DB 62 RWLRRY 67

RESULT 9
BADI7451 PRELIMINARY; PRT; 100 AA.
ID BADI7451;
AC BADI7451;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSUNBB0046012.18.
GN OSUNBB0046012.18.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP005643; BADI7451.1; -.
DR EMBL; AP005644; BADI7502.1; -.
KW Hypothetical protein.
SQ SEQUENCE 100 AA; 10960 MW; CC21DOB5AFBD0E4D CRC64;

Query Match 30.0%; Score 6; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLRR 15
DB 19 WRWLRR 24

RESULT 10
BADI7502 PRELIMINARY; PRT; 100 AA.
ID BADI7502;
AC BADI7502;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNB0056122.58.
GN OSJNB0056122.58.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP005643; BADI7451.1; -.
DR EMBL; AP005644; BADI7502.1; -.
KW Hypothetical protein.
SQ SEQUENCE 100 AA; 10960 MW; CC21DOB5AFBD0E4D CRC64;

Query Match 30.0%; Score 6; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLRR 15
DB 19 WRWLRR 24

RESULT 11
Q6Z4J8 PRELIMINARY; PRT; 101 AA.
ID Q6Z4J8;
AC Q6Z4J8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
```

```

DE Hypothetical protein OSJNBb0002J01.11.
GN Name=OSJNBb0002J01.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarthroideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF005172; BAC83842.1; -.
KW Hypothetical protein.
SQ SEQUENCE 101 AA; 10261 MW; 238F8F70BF4D8092 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRLRR 15
|||
DB 66 WRLRR 71

RESULT 12
BAC83842 PRELIMINARY; PRT; 101 AA.
AC BAC83842;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNBb0002J01.11.
GN OSJNBb0002J01.11.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarthroideae; Oryzeae; Oryza; Oryza sativa.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RC Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GAS) genomic DNA, chromosome 7, BAC
clone:OSJNBb0002J01.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF005172; BAC83842.1; -.
KW Hypothetical protein.
SQ SEQUENCE 101 AA; 10261 MW; 238F8F70BF4D8092 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRLRR 15
|||
DB 66 WRLRR 71

RESULT 13
QBHG83 PRELIMINARY; PRT; 126 AA.
AC QBHG83;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
OS Dennyus hirsundinis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Phthiraptera; Amblycera; Menoponidae; Dennyus.
OX NCBI_TaxID=60508;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=22454480; PubMed=12565034;
RA Johnson K.P., Cruickshank R.H., Adams R.J., Smith V.S., Page R.D.,
RA Clayton D.H.;
RT "Dramatically elevated rate of mitochondrial substitution in lice
(RT (Insecta: Phthiraptera).";
RL Mol. Phylogenet. Evol. 26:231-242(2003).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. CO I is the
CC catalytic subunit of the enzyme. Electrons originating in
CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the bimetallic center formed by heme A3
CC and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AF545694; AAO11921.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000893; COX1.
DR Pfam; PF00115; COX1; 1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON_TER 1 126
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 13594 MW; 3CA5B95D61A5487F CRC64;

Query Match 30.0%; Score 6; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
|||
DB 97 SLWALG 102

RESULT 14
Q7YBD3 PRELIMINARY; PRT; 127 AA.
AC Q7YBD3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Echinmpteryx hageni.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Psocoptera; Trogiomorpha; Lepidopsocidae;
OC Echinmpteryx.
OX NCBI_TaxID=209971;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnson K.P., Mockford B.L.;
RT "Molecular systematics of Psocomorpha (Psocoptera).";
RL Syst. Entomol. 28:409-416(2003).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. CO I is the
CC catalytic subunit of the enzyme. Electrons originating in
CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the bimetallic center formed by heme A3
CC and copper B (By similarity).

```

GO; GO:0016491; F:oxygenase activity; IEA.
GO; GO:0006118; P:electron transport; IEA.
GO; GO:0006810; P:transport; IEA.
InterPro; IPR000883; COX1.
Pfam; PF00115; COX1; 1.
PRINTS; PR01165; CYCOXIDASE1.
Copper; Electron transport; Heme; Inner membrane; Membrane;
Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
Transport.
KW FT NON TER
FT NON TER
FT NON TER
SQ SEQUENCE
Query Match
Best Local Similarity
Matches
QY
DB
RESULT 16
Q8SI37 PRELIMINARY; PRT; 134 AA.
ID Q8SI37
AC Q8SI37
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Ozopemon brownel.
OZ Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Scolytidae; Ozopemon.
OX NCBI_TaxID=124017;
RN [1]
SEQUENCE FROM N.A.
RA Jordal B.H., Beaver R.A., Normark B.B., Farrell B.D.;
RP "Extraordinary sex ratios and the evolution of male neoteny in sib-
RT mating Ozopemon beetles";
RL Biol. J. Linn. Soc. Lond. 75:353-360(2002).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. CO I is the
CC catalytic subunit of the enzyme. Electrons originating in
CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the bimetallic center formed by heme A3
CC and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocyanide c + O(2) = 4 ferricyanide
CC c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AF438510; AAC08274.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F:oxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000883; COX1.
Pfam; PF00115; COX1; 1.
Copper; Electron transport; Heme; Inner membrane; Membrane;
Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
Transport.
KW FT NON TER
FT NON TER
FT NON TER
SQ SEQUENCE

```

Query Match          30.0%; Score 6; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
DB 97 SLWALG 102

RESULT 17
Q8VWR7 PRELIMINARY; PRT; 143 AA.
AC Q8VWR7;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Transporter-like protein (Fragment).
OS Narcissus pseudonarcissus (Daffodil).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;
OC Narcissus.
OX NCBI_TaxID=39639;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=4-day old flower tepal;
RA Hunter D.A., Steele B.C., Reid M.S.;
RT "Identification of genes associated with perianth senescence in
RT daffodil (Narcissus pseudonarcissus L. 'Dutch Master').";
RL Plant Sci. 163:13-21(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AF462230; AAC69393.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar_tr; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 143
FT NON_TER 143
SQ SEQUENCE 143 AA; 15957 MW; E5AB86F1F30879F CRC64;

Query Match          30.0%; Score 6; DB 2; Length 143;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LGWRWL 13
DB 11 LGWRWL 16

RESULT 18
Q84010 PRELIMINARY; PRT; 162 AA.
AC Q84010;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Pseudomonas chlororaphis (Pseudomonas aureofaciens).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=333;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim J.J., Lawrence J.E., Kluepfel D.A.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY163858; AAP30022.1; -.
KW Hypothetical protein.
SQ SEQUENCE 162 AA; 19205 MW; 6AD69FCF5A3689DA CRC64;

Query Match          30.0%; Score 6; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 69;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRYG 17
DB 124 WLRRYG 129

RESULT 19
Q73NA1 PRELIMINARY; PRT; 170 AA.
ID Q73NA1;
AC Q73NA1;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Smr domain protein.
GN OrderedlocusNames=TDE1254;
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidsen T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,
RA Gebregorgis E., Geer K., Tsegaye G., Malek J.A., Ayodeji B.,
RA Shatsman P., McLeod M.P., Smajs D., Howell J.K., Pal S., Amin A.,
RA Vashisth P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
RA Weinstein G.M., Norris S.J., Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
RT with other spirochete genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
DR EMBL; AE017250; AAS11772.1; -.
DR TIGR; TDE1254; -.
DR InterPro; IPR002625; Smr/MutS2_C.
DR Pfam; PF01713; Smr; 1.
DR SMART; SM00463; SMR; 1.
KW Complete proteome.
SQ SEQUENCE 170 AA; 19438 MW; 9A29EB2307D04319 CRC64;

Query Match          30.0%; Score 6; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRYG 17
DB 54 WLRRYG 59

RESULT 20
AAS11772 PRELIMINARY; PRT; 170 AA.
ID AAS11772;
AC AAS11772;
DT 02-MAR-2004 (TREMBLrel. 27, Created)
DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
DT 04-MAY-2004 (TREMBLrel. 27, Last annotation update)
DE Smr domain protein.
GN TDE1254.
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidsen T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J., Durkin S.A.,
RA Daugherty S.C., Shetty J., Shvartsbeyn A., Gebregorgis E., Geer K.,
RA Tsegaye G., Malek J.A., Ayodeji B., Shatsman S., McLeod M.P.,
RA Smajs D., Howell J.K., Pal S., Amin A., Vashisth P., McNeill T.Z.,

```

RA Xiang Q., Sodergren E., Baca E., Weinstock G.M., Norris S.J.,
 RA Fraser C.M., Paulsen I.T.,
 RT "Comparison of the genome of the oral pathogen *Treponema denticola*
 RT with other spirochete genomes.",
 RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
 DR EMBL; AE017250; AAS11772.1; -.
 DR TIGR; TDE1254; -.
 SQ SEQUENCE 170 AA; 19438 MW; 9A29EB2307D04319 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 170;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 WRRYRG 17
 Db 54 WRRYRG 59
 |||||

RESULT 21
 QSHAT4
 ID QSHAT4 PRELIMINARY; PRT; 171 AA.
 AC QSHAT4;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).
 GN Name=COI;
 OS Psacothaea hillaris yonaguniana.
 OG Mitochondrion.
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 CC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini;
 CC Psacothaea.
 CC NCBI_TaxID=192788;
 CC [1]
 RN SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;
 RT "Phylogenetic analysis of the phoretic association between
 RT *Bursaphelenchus conicaudatus* (Nematoda: Aphelenchoididae) and
 RT *Psacothaea hillaris* (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the bimetallic center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AB083741; BAC44945.1; -.
 DR EMBL; AB083742; BAC44946.1; -.
 DR EMBL; AB083743; BAC44947.1; -.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1, 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 171

SQ SEQUENCE 171 AA; 19210 MW; FB9848BF92FC0DEC CRC64;

Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLWALG 9
 Db 15 SLWALG 20
 |||||

RESULT 22
 QSHAT5
 ID QSHAT5 PRELIMINARY; PRT; 171 AA.
 AC QSHAT5;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).
 GN Name=COI;
 OS Psacothaea hillaris miyakoana.
 OG Mitochondrion.
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 CC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini;
 CC Psacothaea.
 CC NCBI_TaxID=192790;
 CC [1]
 RN SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;
 RT "Phylogenetic analysis of the phoretic association between
 RT *Bursaphelenchus conicaudatus* (Nematoda: Aphelenchoididae) and
 RT *Psacothaea hillaris* (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the bimetallic center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AB083750; BAC44954.1; -.
 DR EMBL; AB083751; BAC44955.1; -.
 DR EMBL; AB083752; BAC44956.1; -.
 DR EMBL; AB083753; BAC44957.1; -.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1, 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 171

SQ SEQUENCE 171 AA; 19210 MW; FB9848BF92FC0DEC CRC64;

Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
Db 15 SLWALG 20

RESULT 23

```

QBHAT6
ID Q8HAT6 PRELIMINARY; PRT; 171 AA.
AC Q8HAT6;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit 1 (Fragment).
GN Name=COL;
OS Psacothaea hilaris maculata.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini;
OC Psacothaea.
OX NCBI_TaxID=192793;
RN [1]_TaxID=192793;
RP SEQUENCE FROM N.A.
RA Kanzaki N., Futai K.;
RT "Phylogenetic analysis of the phoretic association between
RT Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and
RT Psacothaea hilaris (Coleoptera: Cerambycidae).";
RL Nematology 4:759-771(2002).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. CO I is the
CC catalytic subunit of the enzyme. Electrons originating in
CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the binuclear center formed by heme A3
CC and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AB083760; BAC44964.1; -.
DR EMBL; AB083761; BAC44965.1; -.
DR EMBL; AB083762; BAC44966.1; -.
DR EMBL; AB083763; BAC44967.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON_TER 1 1
FT NON_TER 171 171
SQ SEQUENCE 171 AA; 19210 MW; FB9848BF92FC0DEC CRC64;

```

Query Match 30.0%; Score 6; DB 2; Length 171;
Best Local Similarity 100.0%; Pred.No.72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
Db 15 SLWALG 20

RESULT 24

```

QBHAT7
ID Q8HAT7 PRELIMINARY; PRT; 171 AA.
AC Q8HAT7;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit 1 (Fragment).
GN Name=COL;
OS Psacothaea hilaris macronotata.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini;
OC Psacothaea.
OX NCBI_TaxID=192794;
RN [1]_TaxID=192794;
RP SEQUENCE FROM N.A.
RA Kanzaki N., Futai K.;
RT "Phylogenetic analysis of the phoretic association between
RT Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and
RT Psacothaea hilaris (Coleoptera: Cerambycidae).";
RL Nematology 4:759-771(2002).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. CO I is the
CC catalytic subunit of the enzyme. Electrons originating in
CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the binuclear center formed by heme A3
CC and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AB083764; BAC44968.1; -.
DR EMBL; AB083765; BAC44969.1; -.
DR EMBL; AB083766; BAC44970.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON_TER 1 1
FT NON_TER 171 171
SQ SEQUENCE 171 AA; 19210 MW; FB9848BF92FC0DEC CRC64;

```

Query Match 30.0%; Score 6; DB 2; Length 171;
Best Local Similarity 100.0%; Pred.No.72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
Db 15 SLWALG 20

RESULT 25

```

QBHAT8
ID Q8HAT8 PRELIMINARY; PRT; 171 AA.
AC Q8HAT8;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit 1 (Fragment).

```

GN Name=COL;
 OS Psacothaea hilaris ishigakiana.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini;
 OC Psacothaea.
 OX NCBI_TaxID=192789;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;
 RT "Phylogenetic analysis of the phoretic association between
 RT Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and
 RT Psacothaea hilaris (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the bimetallic center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AB083744; BAC44948.1; -;
 DR EMBL; AB083745; BAC44949.1; -;
 DR EMBL; AB083746; BAC44950.1; -;
 DR EMBL; AB083747; BAC44951.1; -;
 DR EMBL; AB083748; BAC44952.1; -;
 DR EMBL; AB083749; BAC44953.1; -;
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON TER 1 1
 FT NON TER 171 171
 SQ SEQUENCE 171 AA; 19210 MW; FB9848BF92FCODEC CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWALG 9 PRT; 171 AA.
 DB 15 SLWALG 20
 RESULT 26
 Q8HAT9
 ID Q8HAT9 PRELIMINARY; PRT; 171 AA.
 AC Q8HAT9;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).
 GN Name=COL;
 OS Psacothaea hilaris intermedia.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini;
 OC Psacothaea.
 OX NCBI_TaxID=192792;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;
 RT "Phylogenetic analysis of the phoretic association between
 RT Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and
 RT Psacothaea hilaris (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the bimetallic center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AB083756; BAC44960.1; -;
 DR EMBL; AB083757; BAC44961.1; -;
 DR EMBL; AB083759; BAC44963.1; -;
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON TER 1 1
 FT NON TER 171 171
 SQ SEQUENCE 171 AA; 19210 MW; FB9848BF92FCODEC CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWALG 9 PRT; 171 AA.
 DB 15 SLWALG 20
 RESULT 27
 Q8HAU0
 ID Q8HAU0 PRELIMINARY; PRT; 171 AA.
 AC Q8HAU0;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).
 GN Name=COL;
 OS Psacothaea hilaris insularis.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini;
 OC Psacothaea.
 OX NCBI_TaxID=192795;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;

RT "Phylogenetic analysis of the phoretic association between
 RT Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and
 RT Psacothaea hilaris (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the binuclear center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AB083770; BAC44974.1; -.
 DR EMBL; AB083771; BAC44975.1; -.
 DR EMBL; AB083772; BAC44976.1; -.
 DR EMBL; AB083773; BAC44977.1; -.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016023; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000893; COX1.
 DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 1
 FT NON_TER 171 171
 SQ SEQUENCE 171 AA; 19210 MW; FB9848BF92FCODEC CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWALG 9
 Db 15 SLWALG 20
 RESULT 28
 Q8HAU1 PRELIMINARY; PRT; 171 AA.
 ID Q8HAU1
 AC Q8HAU1;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).
 OS Name=COL;
 GN Psacothaea hilaris
 OS Psacothaea hilaris
 OG Mitochondrion.
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 CC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini;
 CC Psacothaea.
 CC NCBI_TaxID=192796;
 RN [1]_TaxID=192796;
 RP SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;
 RT "Phylogenetic analysis of the phoretic association between
 RT Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and
 RT Psacothaea hilaris (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the binuclear center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.

CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the binuclear center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AB083776; BAC44980.1; -.
 DR EMBL; AB083777; BAC44981.1; -.
 DR EMBL; AB083778; BAC44982.1; -.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016023; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000893; COX1.
 DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 1
 FT NON_TER 171 171
 SQ SEQUENCE 171 AA; 19210 MW; FB9848BF92FCODEC CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWALG 9
 Db 15 SLWALG 20
 RESULT 29
 Q8HAU2 PRELIMINARY; PRT; 171 AA.
 ID Q8HAU2
 AC Q8HAU2;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).
 OS Name=COL;
 GN Psacothaea hilaris
 OS Psacothaea hilaris
 OG Mitochondrion.
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 CC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini;
 CC Psacothaea.
 CC NCBI_TaxID=192794;
 RN [1]_TaxID=192794;
 RP SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;
 RT "Phylogenetic analysis of the phoretic association between
 RT Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and
 RT Psacothaea hilaris (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the binuclear center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.

CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AB083767; BAC44971.1; -.
 DR EMBL; AB083769; BAC44973.1; -.
 DR EMBL; AB083768; BAC44972.1; -.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.
 KW FT NON_TER 1 171
 FT NON_TER 171 171
 SQ SEQUENCE 171 AA; 19229 MW; FB9848BED7FC0DEC CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWALG 9
 DB 15 SLWALG 20
 RESULT 30
 Q8HY9 PRELIMINARY; PRT; 171 AA.
 AC Q8HY9; 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).
 GN Name=COL;
 OS Psacothaea hilaris hilaris.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia; Phycophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini; Psacothaea.
 OC NCBI_TaxID=192796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;
 RT "Phylogenetic analysis of the phoretic association between Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and Psacothaea hilaris (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme. Electrons originating in catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the bimetallic center formed by heme A3 and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferrocycytochrome c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family. EMBL; AB083774; BAC44978.1; -.
 DR EMBL; AB083774; BAC44978.1; -.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0004129; F:oxidoreductase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.

DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.
 KW FT NON_TER 1 171
 FT NON_TER 171 171
 SQ SEQUENCE 171 AA; 19223 MW; 5596A84162E0C0FO CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWALG 9
 DB 15 SLWALG 20
 RESULT 31
 Q8HI88 PRELIMINARY; PRT; 171 AA.
 AC Q8HI88; 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).
 GN Name=COL;
 OS Psacothaea hilaris miyakejimana.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia; Phycophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agniini; Psacothaea.
 OC NCBI_TaxID=192797;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;
 RT "Phylogenetic analysis of the phoretic association between Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and Psacothaea hilaris (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the bimetallic center formed by heme A3 and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferrocycytochrome c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family. EMBL; AB083775; BAC44979.1; -.
 DR EMBL; AB083775; BAC44979.1; -.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0004129; F:oxidoreductase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.

DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 FT Transport.
 FT NON_TER 1 1
 FT NON_TER 171 171
 SQ SEQUENCE 171 AA; 19210 MW; FB9848BF92FCODEC CRC64;

Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 Db 15 SLWALG 20
 |||||

RESULT 32

Q8HI89 PRELIMINARY; PRT; 171 AA.
 AC Q8HI89;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).

OS Name=COL;
 GS Psacothaea hilaris intermedia.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Artropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agnini;
 OC Psacothaea.
 OX NCBI_TaxID=192792;
 RN [1]

RP SEQUENCE FROM N.A.

RA Kanazaki N., Futai K.;

RT "Phylogenetic analysis of the phoretic association between
 Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and
 Psacothaea hilaris (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).

CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the bimetallic center formed by heme A3
 CC and copper B (By similarity).

CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.

CC -!- PATHWAY: Respiratory chain; terminal step.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial

CC inner membrane (By similarity).

CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.

DR EMBL; AB083758; BAC44962.1; -.

DR GO; GO:0019866; C:inner membrane; IEA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

DR GO; GO:0005739; C:mitochondrion; IEA.

DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000883; COX1.

DR Pfam; PF01115; COX1; 1.

DR PRINTS; PR01165; CYCOXIDASE1.

KW Copper; Electron transport; Heme; Inner membrane; Membrane;

KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;

KW Transport.

FT NON_TER 1 1

FT NON_TER 171 171

SQ SEQUENCE 171 AA; 19240 MW; FB9848A5385D1CEC CRC64;

Query Match

30.0%; Score 6; DB 2; Length 171;

Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 33

Q8HI90 PRELIMINARY; PRT; 171 AA.
 AC Q8HI90;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).

OS Name=COL;

GS Psacothaea hilaris intermedia.

OG Mitochondrion.

OC Eukaryota; Metazoa; Artropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;

OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agnini;

OC Psacothaea.

OX NCBI_TaxID=192792;

RN [1]

RP SEQUENCE FROM N.A.

RA Kanazaki N., Futai K.;

RT "Phylogenetic analysis of the phoretic association between

Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and

Psacothaea hilaris (Coleoptera: Cerambycidae).";

RL Nematology 4:759-771(2002).

CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the bimetallic center formed by heme A3
 CC and copper B (By similarity).

CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.

CC -!- PATHWAY: Respiratory chain; terminal step.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).

CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.

DR EMBL; AB083758; BAC44959.1; -.

DR GO; GO:0019866; C:inner membrane; IEA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

DR GO; GO:0005739; C:mitochondrion; IEA.

DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000883; COX1.

DR Pfam; PF01115; COX1; 1.

DR PRINTS; PR01165; CYCOXIDASE1.

KW Copper; Electron transport; Heme; Inner membrane; Membrane;

KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;

KW Transport.

FT NON_TER 1 1

FT NON_TER 171 171

SQ SEQUENCE 171 AA; 19186 MW; B7389CEB97B91A51 CRC64;

Query Match

30.0%; Score 6; DB 2; Length 171;

Best Local Similarity 100.0%; Pred. No. 72;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9

Db 15 SLWALG 20

|||||

RESULT 34

Q8HI91
 ID Q8HI91 PRELIMINARY; PRT; 171 AA.
 AC Q8HI91;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).
 GN Name=COL;
 OS Psacotha hilaris tenebrosa.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Agnini;
 OC Psacotha.
 OC NCBI_TaxID=192791;
 OX NCBI_TaxID=192791;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;
 RT "Phylogenetic analysis of the phoretic association between
 RT Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and
 RT Psacotha hilaris (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the bimetallic center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AB083754; BAC4958.1; -.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 171
 FT NON_TER 171 171
 SQ SEQUENCE 171 AA; 19229 MW; 1B8C19FFC6FAA72A CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred.No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWALG 9
 Db 15 SLWALG 20
 RESULT 35
 Q8HI92
 ID Q8HI92 PRELIMINARY; PRT; 171 AA.
 AC Q8HI92;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Cytochrome oxidase subunit 1 (Fragment).
 GN Name=COL;
 OS Monochamus alternatus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Symphyla; Scutigereillidae;
 OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Symphyla; Scutigereillidae;
 OC Hansenella.
 OC NCBI_TaxID=134582;
 OX NCBI_TaxID=134582;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thampapillai G.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory

OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Monochamini;
 OC Monochamus.
 OX NCBI_TaxID=192382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanzaki N., Futai K.;
 RT "Phylogenetic analysis of the phoretic association between
 RT Bursaphelenchus conicaudatus (Nematoda: Aphelenchoididae) and
 RT Psacotha hilaris (Coleoptera: Cerambycidae).";
 RL Nematology 4:759-771(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the bimetallic center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AB083740; BAC44944.1; -.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 171
 FT NON_TER 171 171
 SQ SEQUENCE 171 AA; 19196 MW; 269D5136AE243DFF CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 171;
 Best Local Similarity 100.0%; Pred.No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWALG 9
 Db 15 SLWALG 20
 RESULT 36
 Q955W8
 ID Q955W8 PRELIMINARY; PRT; 184 AA.
 AC Q955W8;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Cytochrome oxidase subunit I (fragment).
 GN Name=COX1;
 OS Hanseniella sp. EW-2000.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Symphyla; Scutigereillidae;
 OC Hanseniella.
 OC NCBI_TaxID=134582;
 OX NCBI_TaxID=134582;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thampapillai G.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory

chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the bimetallic center formed by heme A3 and copper B (By similarity).

CC -!- CATALYTIC ACTIVITY: 4 ferrocycochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.

CC -!- PATHWAY: Respiratory chain; terminal step.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).

CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.

DR EMBL; AF218271; AAK94875.1; -.

DR GO; GO:0019866; C:inner membrane; IEA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

DR GO; GO:0005739; C:mitochondrion; IEA.

DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000883; COX1.

DR Pfam; PF00115; COX1.1.

DR PRINTS; PR01165; CYCOXIDASE1.

KW Copper; Electron transport; Heme; Inner membrane; Membrane;

KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;

KW Transport.

FT NON_TER 1 184

FT TER 184 184

SQ SEQUENCE 184 AA; 20155 MW; 24861A9C2CDA9FB0 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 184;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
|||||

DB 95 SLWALG 100

RESULT 37

Q9PEC9 Q9PEC9 PRELIMINARY; PRT; 191 AA.

AC Q9PEC9;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=Xf1099;

OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xylella.

OX NCBI_TaxID=2371;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=9a5c;

RX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,

RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,

RA Bueno M.R.F., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,

RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,

RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,

RA Krieger J.B., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.B.N., Martins E.M.F., Marino C.L.,

RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,

RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,

RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,

RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,

RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

RA de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,

RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,

RA Vallada H., Van Sluys M.A., Verjowski-Almeida S., Vettore A.L.,

RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;

RT "The genome sequence of the plant pathogen Xylella fastidiosa.";

RL Nature 406:151-159(2000).

DR EMBL; AE003945; AAF83909.1; -.

DR PIR; F82725; F82725.

DR InterPro; IPR008934; AcPase_VanPerase.

DR InterPro; IPR000326; Pesterase_PA_PTP.

DR Pfam; PF01569; PAP2; 1.

DR Complete proteome; Hypothetical protein.

SQ SEQUENCE 191 AA; 20144 MW; 4CD570C46B6CFD11 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
|||||

DB 94 SLWALG 99

RESULT 38

Q8M7A7 Q8M7A7 PRELIMINARY; PRT; 213 AA.

AC Q8M7A7;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Cytochrome oxidase subunit I (Fragment).

GN Name=COI;

OS Lycopodium carinatus.

OG Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;

OC Polyphaga; Curculionidae; Molytinae; Molytini; Lycopodium.

OX NCBI_TaxID=169099;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21651553; PubMed=11794778;

RA Treweek S.A., Wallis G.P.;

RT "Bridging the 'beech-gap': New Zealand invertebrate phylogeography

RT implicates Pleistocene glaciation and Pliocene isolation.";

RL Evolution 55:2170-2180(2001)

CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the bimetallic center formed by heme A3 and copper B (By similarity).

CC -!- CATALYTIC ACTIVITY: 4 ferrocycochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.

CC -!- PATHWAY: Respiratory chain; terminal step.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).

CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.

DR EMBL; AY042350; AAL18183.1; -.

DR GO; GO:0019866; C:inner membrane; IEA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

DR GO; GO:0005739; C:mitochondrion; IEA.

DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR GO; GO:0006810; P:transport; IEA.

```

DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON TER 1 1
FT NON TER 213 213
SQ SEQUENCE 213 AA; 23647 MW; B86440BEA5FD6AC8 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLWALG 9
Db 89 SLWALG 94
|||||

RESULT 39
Q6F7R0 PRELIMINARY; PRT; 246 AA.
AC Q6F7R0;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=ACIAD3221;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP SEQUENCE FROM N.A.
RA Barbe V., Vallet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 0:0-0(2004).
DR EMBL; CR543861; CAG69905.1; -.
DR InterPro; IPR006220; Anth_synthII.
DR Pfam; PF00117; GATase_1.
DR PRINTS; PR00097; ANTSNTHASEII.
DR PRINTS; PR00096; GATASE.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 246 AA; 27971 MW; 99A44BC1DD082989 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 246;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RWLRRY 16
Db 86 RWLRRY 91
|||||

RESULT 40
Q9EWC2 PRELIMINARY; PRT; 256 AA.
AC Q9EWC2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Thioesterase.
GN Name=cant;
OS Streptomyces griseus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=IMRU 3570;
RX MEDLINE=21642576; PubMed=11782498;
RA Campelo A.B.; Gil J.A.;
RT "The candidin gene cluster from Streptomyces griseus IMRU 3570.";
RL Microbiology 148:51-59(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IMRU 3570;
RA Campelo A.B.;
RL Thesis (2000), Department of Microbiologia, Universidad de Leon, Leon, Spain.
RL EMBL; AJ300302; CAC22116.1; -.
DR HSSP; Q08787; 1JMK.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR000379; Ser_estrs.
DR InterPro; IPR01031; Thioesterase.
DR Pfam; PF00975; Thioesterase; 1.
SQ SEQUENCE 256 AA; 28080 MW; F65C9F94E11857E5 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 256;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RWLRRY 16
Db 10 RWLRRY 15
|||||

RESULT 41
Q9HYC6 PRELIMINARY; PRT; 266 AA.
AC Q9HYC6;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=PA3483;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.D., Huftagie W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004769; AAG06871.1; -.
DR PIR; D83210; D83210.
DR InterPro; IPR003006; IG_MHC.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 266 AA; 28619 MW; EB241FC27F0660F7 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 WRWLR 15
Db 254 WRWLR 259
|||||

RESULT 42

```


Q6VZL3
ID Q6VZL3 PRELIMINARY; PRT; 270 AA.
AC Q6VZL3;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Mantodea sp. JD-2003.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Mantodea.
OX NCBI_TaxID=253121;
RN [1]
RP SEQUENCE FROM N.A.
RA Klass K.-D., Picker M.D., Damgaard J., van Noort S., Tojo K.;
RT "The Taxonomy, Genitalic Morphology, and Phylogenetic Relationships of
South African Mantophasmatodea (Insecta).";
RL Entomol. Abh. 61:3-67(2003).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. CO I is the
catalytic subunit of the enzyme. Electrons originating in
cytochrome c are transferred via the copper A center of subunit 2
and heme A of subunit 1 to the binuclear center formed by heme A3
and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AY318914; AAR17105.1; -;
DR GO; GO:0005739; C:mitochondrion; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON_TER 1 1
FT NON_TER 270 270
SQ SEQUENCE 270 AA; 30323 MW; 7A57C40E6ED5FDD9 CRC64;
Query Match 30.0%; Score 6; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 SLWALG 9
Db 98 SLWALG 103

RESULT 43
AAR17105 PRELIMINARY; PRT; 270 AA.
ID AAR17105;
AC AAR17105;
DT 02-NAR-2004 (TReMBLrel. 27, Created)
DT 02-NAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-NAR-2004 (TReMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN COI.
OS Mantodea sp. JD-2003.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Mantodea.
OX NCBI_TaxID=253121;
RN [1]
RP SEQUENCE FROM N.A.
RA Klass K.-D., Picker M.D., Damgaard J., van Noort S., Tojo K.;
RT "The Taxonomy, Genitalic Morphology, and Phylogenetic Relationships of
South African Mantophasmatodea (Insecta).";
RL Entomol. Abh. 61:3-67(2003).

DR EMBL; AY318914; AAR17105.1; -;
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 270 270
SQ SEQUENCE 270 AA; 30323 MW; 7A57C40E6ED5FDD9 CRC64;
Query Match 30.0%; Score 6; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 SLWALG 9
Db 98 SLWALG 103

RESULT 44
Q8TVN5 PRELIMINARY; PRT; 277 AA.
ID Q8TVN5;
AC Q8TVN5;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 26, Last annotation update)
DE Uncharacterized protein specific for M.kandleri, MK-45 family.
GN OrderedLocusNames=MK1353;
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaeva K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.D., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozyavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AE010428; AAM02566.1; -;
SQ SEQUENCE 277 AA; 30678 MW; 872BBD49BB95F3A CRC64;
Query Match 30.0%; Score 6; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 WRWLR 15
Db 272 WRWLR 277

RESULT 45
Q6W5Q5 PRELIMINARY; PRT; 285 AA.
ID Q6W5Q5;
AC Q6W5Q5;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE FscTE.
GN Name=fscTE;
OS Streptomyces sp. FR-008.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=206662;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen S., Huang X., Zhou X., Bai L., He J., Jeong K.J., Lee S.Y.,
RA Deng Z.;
RT "Organizational and mutational analysis of a complete FR-
008/candidin gene cluster encoding a structurally related polyene
complex.";
RT

```

RL Chem. Biol. 10:1065-1076(2003).
DR ENBL: AY310323; AAQ82559.1; -.
DR InterPro: IPR000379; Ser_estrs.
DR InterPro: IPR001031; Thioesterase.
DR Pfam: PF00975; Thioesterase; 1.
SQ SEQUENCE 285 AA; 31691 MW; FD3CABEB2ABA7A92 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RWLRRY 16
Db 39 RWLRRY 44

RESULT 46
AAQ82559 PRELIMINARY; PRT; 285 AA.
AC AAQ82559;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE FSCTE.
GN Streptomyces sp. FR-008.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=206662;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14652074;
RA Chen S., Huang X., Zhou X., Bai L., He J., Jeong K.J., Lee S.Y.,
RA Deng Z.;
RT "Organizational and Mutational Analysis of a Complete FR-
RT 008/Candidin Gene Cluster Encoding a Structurally Related Polyene
RT Complex.";
RL Chem. Biol. 10:1065-1076(2003).
DR ENBL: AY310323; AAQ82559.1; -.
SQ SEQUENCE 285 AA; 31691 MW; FD3CABEB2ABA7A92 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RWLRRY 16
Db 39 RWLRRY 44

RESULT 47
Q7U6Q5 PRELIMINARY; PRT; 288 AA.
AC Q7U6Q5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ABC transporter for possibly for trehalose/maltose, membrane component
DE precursor.
GN OrderedLocustNames=SYNW1283;
OS Synecchococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synecchococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahmsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synecchococcus.";
RL Nature 424:1037-1042(2003).
CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across

the membrane (By similarity).
-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
-!- SIMILARITY: Belongs to the binding-protein-dependent transport
system permease family.
EMBL: BX569692; CAB07798.1; -.
GO: GO:0016020; C:membrane; IEA.
GO: GO:0005215; F:transporter activity; IEA.
GO: GO:0006810; P:transport; IEA.
InterPro: IPR000515; BPD_transp.
Pfam: PF00528; BPD_transp_1; 1.
PROSITE: PS00928; ABC_TM1; 1.
KW Complete proteome; Signal; Transmembrane; Transport.
FT SIGNAL 288 32 Potential.
FT SIGNAL 1 32 Potential.
SQ SEQUENCE 288 AA; 31458 MW; 0EF08A3C854BD843 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 ALGWRW 12
Db 112 ALGWRW 117

RESULT 48
Q7V7P5 PRELIMINARY; PRT; 288 AA.
AC Q7V7P5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ABC transporter for sugars, membrane component precursor.
GN OrderedLocustNames=PMT0693;
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;
RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
the membrane (By similarity).
-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
-!- SIMILARITY: Belongs to the binding-protein-dependent transport
system permease family.
EMBL: BX572097; CAE20868.1; -.
GO: GO:0016020; C:membrane; IEA.
GO: GO:0005215; F:transporter activity; IEA.
GO: GO:0006810; P:transport; IEA.
InterPro: IPR000515; BPD_transp.
Pfam: PF00528; BPD_transp_1; 1.
PROSITE: PS00928; ABC_TM1; 1.
KW Complete proteome; Signal; Transmembrane; Transport.
FT SIGNAL 288 32 Potential.
FT SIGNAL 1 32 Potential.
SQ SEQUENCE 288 AA; 32176 MW; 506FA8DFD2D2E4825 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 ALGWRW 12
Db 112 ALGWRW 117

```


RX MEDLINE=21596578; PubMed=11761062;
RA Jordal B.D., Sequeira A.S., O'Weara B.C., Normark B.B., Chung J.H.,
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL Evolution 55:2011-2027(2001).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. CO I is the
catalytic subunit of the enzyme. Electrons originating in
cytochrome c are transferred via the copper A center of subunit 2
and heme A of subunit 1 to the bimetallic center formed by heme A3
and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL: AF375329; AA057415.1; -.
DR GO: GO:0019866; C:inner membrane; IEA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000883; COX1.
DR Pfam: PF00115; COX1; 1.
DR PRINTS: PS00077; COX1; 1.
DR PROSITE: PS00077; COX1; 1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON_TER 327 327
FT FT
SQ SEQUENCE 327 AA; 35766 MW; C81C79B2C01B47A9 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLWALG 9
Db 242 SLWALG 247
|||||
RESULT 53
Q8SI27 PRELIMINARY; PRT; 330 AA.
AC Q8SI27;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Samponius dampfi.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Scolytidae; Samponius.
OX NCBI_TaxID=124019;
RN [1]
RP SEQUENCE FROM N.A.
RA Jordal B.H., Beaver R.A., Normark B.B., Farrell B.D.;
RT "Extraordinary sex ratios and the evolution of male neoteny in sib-
mating Ozopemon beetles.";
RL Biol. J. Linn. Soc. Lond. 75:353-360(2002).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. CO I is the
catalytic subunit of the enzyme. Electrons originating in
cytochrome c are transferred via the copper A center of subunit 2
and heme A of subunit 1 to the bimetallic center formed by heme A3
and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL: AF438519; AA08283.1; -.

CC cytochrome c are transferred via the copper A center of subunit 2
and heme A of subunit 1 to the bimetallic center formed by heme A3
and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL: AF438520; AA08284.1; -.
DR GO: GO:0019866; C:inner membrane; IEA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000883; COX1.
DR Pfam: PF00115; COX1; 1.
DR PRINTS: PS00077; COX1; 1.
DR PROSITE: PS00077; COX1; 1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON_TER 330 330
FT NON_TER
SQ SEQUENCE 330 AA; 36047 MW; 77272C147EAF5EE CRC64;

Query Match 30.0%; Score 6; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLWALG 9
Db 242 SLWALG 247
|||||
RESULT 54
Q8SI28 PRELIMINARY; PRT; 330 AA.
AC Q8SI28;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Ambrosiodmus aegir.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Scolytidae; Ambrosiodmus.
OX NCBI_TaxID=123995;
RN [1]
RP SEQUENCE FROM N.A.
RA Jordal B.H., Beaver R.A., Normark B.B., Farrell B.D.;
RT "Extraordinary sex ratios and the evolution of male neoteny in sib-
mating Ozopemon beetles.";
RL Biol. J. Linn. Soc. Lond. 75:353-360(2002).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. CO I is the
catalytic subunit of the enzyme. Electrons originating in
cytochrome c are transferred via the copper A center of subunit 2
and heme A of subunit 1 to the bimetallic center formed by heme A3
and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL: AF438519; AA08283.1; -.

DR GO:0019866; C:inner membrane; IEA.
 DR GO:0016021; C:integral to membrane; IEA.
 DR GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO:0005739; C:mitochondrion; IEA.
 DR GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006118; P:electron transport; IEA.
 DR GO:0006810; P:transport; IEA.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.
 DR PRINTS: PR01165; CYCOXIDASEI.
 DR PROSITE: PS00077; COX1; 1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 Transport.
 FT NON_TER 1 330
 FT NON_TER 330 330
 SQ SEQUENCE 330 AA; 36037 MW; 71F436572CC8F42A CRC64;

Query Match 30.0%; Score 6; DB 2; Length 330;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||
 Db 242 SLWALG 247

RESULT 55

Q8SI30 PRELIMINARY; PRT; 330 AA.
 ID Q8SI30
 AC Q8SI30;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Cytochrome oxidase subunit I (Fragment).
 GN Name=COI;
 OS Dryocetes autographus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Scolytidae; Dryocetes.
 OX NCBI_TaxID=124009;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jordal B.H., Beaver R.A., Normark B.B., Farrell B.D.;
 RT "Extraordinary sex ratios and the evolution of male neoteny in sib-
 mating Ozopemon beetles."
 RL Biol. J. Linn. Soc. Lond. 75:353-360(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 chain that catalyzes the reduction of oxygen to water. Subunits 1-
 3 form the functional core of the enzyme complex. CO I is the
 catalytic subunit of the enzyme. Electrons originating in
 cytochrome c are transferred via the copper A center of subunit 2
 and heme A of subunit 1 to the bimetallic center formed by heme A3
 and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL: AF438517; AA08281.1; -.
 DR GO:0019866; C:inner membrane; IEA.
 DR GO:0016021; C:integral to membrane; IEA.
 DR GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO:0005739; F:cytochrome-c oxidase activity; IEA.
 DR GO:0004129; F:oxidoreductase activity; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006118; P:electron transport; IEA.
 DR GO:0006810; P:transport; IEA.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.

DR PRINTS: PR01165; CYCOXIDASEI.
 DR PROSITE: PS00077; COX1; 1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 Transport.
 FT NON_TER 1 330
 FT NON_TER 330 330
 SQ SEQUENCE 330 AA; 35836 MW; F72DB3B46742FE6E CRC64;

Query Match 30.0%; Score 6; DB 2; Length 330;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||
 Db 242 SLWALG 247

RESULT 56

Q8SI33 PRELIMINARY; PRT; 330 AA.
 ID Q8SI33
 AC Q8SI33;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Cytochrome oxidase subunit I (Fragment).
 GN Name=COI;
 OS Ozopemon browni.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Scolytidae; Ozopemon.
 OX NCBI_TaxID=124017;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jordal B.H., Beaver R.A., Normark B.B., Farrell B.D.;
 RT "Extraordinary sex ratios and the evolution of male neoteny in sib-
 mating Ozopemon beetles."
 RL Biol. J. Linn. Soc. Lond. 75:353-360(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 chain that catalyzes the reduction of oxygen to water. Subunits 1-
 3 form the functional core of the enzyme complex. CO I is the
 catalytic subunit of the enzyme. Electrons originating in
 cytochrome c are transferred via the copper A center of subunit 2
 and heme A of subunit 1 to the bimetallic center formed by heme A3
 and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL: AF438514; AA08279.1; -.
 DR GO:0019866; C:inner membrane; IEA.
 DR GO:0016021; C:integral to membrane; IEA.
 DR GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO:0005739; F:cytochrome-c oxidase activity; IEA.
 DR GO:0004129; F:oxidoreductase activity; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006118; P:electron transport; IEA.
 DR GO:0006810; P:transport; IEA.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.
 DR PRINTS: PR01165; CYCOXIDASEI.
 DR PROSITE: PS00077; COX1; 1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 Transport.
 FT NON_TER 1 330
 FT NON_TER 330 330
 SQ SEQUENCE 330 AA; 36024 MW; 5AA90D0F07E4956 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 330;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLWALG 9
Db 242 SLWALG 247

RESULT 57

Q8SI34 PRELIMINARY; PRT; 330 AA.
AC Q8SI34;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Ozeponon graniceps.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Scolytidae; Coccotrypes.
OX NCBI_TaxID=124000;
RN [1]
RP SEQUENCE FROM N.A.
RA Jordal B.H., Beaver R.A., Normark B.B., Farrell B.D.;
RT "Extraordinary sex ratios and the evolution of male neoteny in sub-
mating Ozeponon beetles.";
RL Biol. J. Linn. Soc. Lond. 75:353-360(2002).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. CO I is the
catalytic subunit of the enzyme. Electrons originating in
cytochrome c are transferred via the copper A center of subunit 2
and heme A of subunit 1 to the bimetallic center formed by heme A3
and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AF438513; C:inner membrane; IEA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON TER 1 1
FT NON TER 330 330
SQ SEQUENCE 330 AA; 36077 MW; 52B22651C6CE73AF CRC64;

Query Match 30.0%; Score 6; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLWALG 9
Db 242 SLWALG 247

RESULT 58

Q8SI41 PRELIMINARY; PRT; 330 AA.
AC Q8SI41;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Ozeponon uniseriatus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Scolytidae; Ozeponon.
OX NCBI_TaxID=190778;
RN [1]
RP SEQUENCE FROM N.A.
RA Jordal B.H., Beaver R.A., Normark B.B., Farrell B.D.;
RT "Extraordinary sex ratios and the evolution of male neoteny in sub-
mating Ozeponon beetles.";
RL Biol. J. Linn. Soc. Lond. 75:353-360(2002).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. CO I is the
catalytic subunit of the enzyme. Electrons originating in
cytochrome c are transferred via the copper A center of subunit 2
and heme A of subunit 1 to the bimetallic center formed by heme A3
and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AF438506; C:inner membrane; IEA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON TER 1 1
FT NON TER 330 330
SQ SEQUENCE 330 AA; 36039 MW; F6675D7AE504A0A4 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLWALG 9
Db 242 SLWALG 247

RESULT 59

Q911W1 PRELIMINARY; PRT; 331 AA.
AC Q911W1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PA2154;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RA MEDLINE=20437337; PubMed=10984043;
 RX Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltz L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RA "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004642; AAG05542.1; -
 DR PIR; A83377; A83377.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 331 AA; 37001 MW; EFBEAB77CF5FC750 CRC64;

 Query Match 30.0%; Score 6; DB 2; Length 331;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 10 WRWLR 15
 Db 18 WRWLR 23

 RESULT 60
 Q9TAS9 PRELIMINARY; PRT; 337 AA.
 AC Q9TAS9;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Cytochrome oxidase I (Fragment).
 OS Cocciotypes advena.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Scolytidae; Cocciotypes.
 OX NCBI_TaxID=105108;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Normark B.B., Jordal B.H., Farrell B.D.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the binuclear center formed by heme A3
 CC and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferrocyclochrome
 CC c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AF187116; AAF04050.1; -
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004123; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 DR PROSITE; PS00077; COX1; 1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;

KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 1
 FT NON_TER 337 337
 SQ SEQUENCE 337 AA; 36768 MW; FD20B47F8BAFF966 CRC64;

 Query Match 30.0%; Score 6; DB 2; Length 337;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 4 SLWALG 9
 Db 247 SLWALG 252

 RESULT 61
 Q6TFC7 PRELIMINARY; PRT; 349 AA.
 AC Q6TFC7;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE LaeA.
 OS Aspergillus fumigatus (Sartorya fumigata).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5085;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15075281;
 RA Bok J.W., Keller N.P.;
 RT "LaeA, a regulator of secondary metabolism in *Aspergillus* spp.";
 RL Eukaryotic Cell 3:527-535(2004).
 DR EMBL; AY422723; AAR01218.1; -
 DR InterPro; IPR000051; SAM_bind.
 SQ SEQUENCE 349 AA; 40076 MW; 0462C78D1A7B82A0 CRC64;

 Query Match 30.0%; Score 6; DB 2; Length 349;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 4 SLWALG 9
 Db 165 SLWALG 170

 RESULT 62
 Q8SEB6 PRELIMINARY; PRT; 349 AA.
 AC Q8SEB6;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Cytochrome oxidase subunit I (Fragment).
 GN Name=COI;
 OS Palame mimetica.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthocini;
 OC Palame.
 OX NCBI_TaxID=185376;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Berkov A.;
 RT "The impact of redefined species limits in Palame (Coleoptera,
 RT Cerambycidae, Lamiinae, Acanthocini) on assessments of host,
 RT seasonal, and stratum specificity.";
 RL Biol. J. Linn. Soc. Lond. 76:195-209(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the
 CC catalytic subunit of the enzyme. Electrons originating in

cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the bimetallic center formed by heme A3 and copper B (By similarity).

-!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.

-!- PATHWAY: Respiratory chain; terminal step.

-!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).

-!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.

EMBL; AF466961; AAL75534.1; -

EMBL; AF466962; AAL75535.1; -

EMBL; AF466963; AAL75536.1; -

EMBL; AF466964; AAL75537.1; -

EMBL; AF466965; AAL75538.1; -

EMBL; AF466966; AAL75539.1; -

EMBL; AF466967; AAL75540.1; -

EMBL; AF466968; AAL75541.1; -

EMBL; AF466969; AAL75542.1; -

EMBL; AF466970; AAL75543.1; -

EMBL; AF466971; AAL75544.1; -

EMBL; AF466972; AAL75545.1; -

EMBL; AF466973; AAL75546.1; -

EMBL; AF466974; AAL75547.1; -

EMBL; AF466975; AAL75548.1; -

EMBL; AF466976; AAL75549.1; -

GO; GO:0019866; C:inner membrane; IEA.

GO; GO:0016021; C:integral to membrane; IEA.

GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

GO; GO:0005739; C:mitochondrion; IEA.

GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.

GO; GO:0016491; F:oxidoreductase activity; IEA.

GO; GO:0006118; P:electron transport; IEA.

GO; GO:0006810; P:transport; IEA.

InterPro; IPR000883; COX1.

Pfam; PF00115; COX1; 1.

PRINTS; PR01165; CYCOXIDASE1.

PROSITE; PS00077; COX1; 1.

Copper; Electron transport; Heme; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.

NON TER 1 1

FT NON TER 349 349

SQ SEQUENCE 349 AA; 38997 MW; 9E01675E32202D9B CRC64;

Query Match 30.0%; Score 6; DB 2; Length 349;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLWALG 9

Db 177 SLWALG 182

RESULT 63

Q8SEEB PRELIMINARY; PRT; 349 AA.

ID Q8SEEB

AC Q8SEEB; 21, Created

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Cytochrome oxidase subunit I (Fragment).

GN Name=COI;

OS Palame mimetica.

OG Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pserygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia; Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthocnini; Palame.

OX NCBI_TaxID=185376;

RN [1]

RP SEQUENCE FROM N.A.

RA Berkov A.;

RT "The impact of redefined species limits in Palame (Coleoptera, Cerambycidae, Lamiinae, Acanthocnini) on assessments of host, seasonal, and stratum specificity."

RL Biol. J. Linn. Soc. Lond. 76:195-209(2002).

CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the bimetallic center formed by heme A3 and copper B (By similarity).

CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome

c + 2 H(2)O.

-!- PATHWAY: Respiratory chain; terminal step.

-!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).

-!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.

EMBL; AF466961; AAL75534.1; -

EMBL; AF466962; AAL75535.1; -

EMBL; AF466963; AAL75536.1; -

EMBL; AF466964; AAL75537.1; -

EMBL; AF466965; AAL75538.1; -

EMBL; AF466966; AAL75539.1; -

EMBL; AF466967; AAL75540.1; -

EMBL; AF466968; AAL75541.1; -

EMBL; AF466969; AAL75542.1; -

EMBL; AF466970; AAL75543.1; -

EMBL; AF466971; AAL75544.1; -

EMBL; AF466972; AAL75545.1; -

EMBL; AF466973; AAL75546.1; -

EMBL; AF466974; AAL75547.1; -

EMBL; AF466975; AAL75548.1; -

EMBL; AF466976; AAL75549.1; -

GO; GO:0019866; C:inner membrane; IEA.

GO; GO:0016021; C:integral to membrane; IEA.

GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

GO; GO:0005739; C:mitochondrion; IEA.

GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.

GO; GO:0016491; F:oxidoreductase activity; IEA.

GO; GO:0006118; P:electron transport; IEA.

GO; GO:0006810; P:transport; IEA.

InterPro; IPR000883; COX1.

Pfam; PF00115; COX1; 1.

PRINTS; PR01165; CYCOXIDASE1.

PROSITE; PS00077; COX1; 1.

Copper; Electron transport; Heme; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.

NON TER 1 1

FT NON TER 349 349

SQ SEQUENCE 349 AA; 38983 MW; 8DB95CDD99B03374 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 349;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLWALG 9

Db 177 SLWALG 182

RESULT 64

Q8SEI6 PRELIMINARY; PRT; 349 AA.

ID Q8SEI6

AC Q8SEI6; 21, Created

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Cytochrome oxidase subunit I (Fragment).

GN Name=COI;

OS Palame crassimana.

OG Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pserygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia; Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthocnini; Palame.

OX NCBI_TaxID=185373;

RN [1]

RP SEQUENCE FROM N.A.

RA Berkov A.;

RT "The impact of redefined species limits in Palame (Coleoptera, Cerambycidae, Lamiinae, Acanthocnini) on assessments of host, seasonal, and stratum specificity."

RL Biol. J. Linn. Soc. Lond. 76:195-209(2002).

CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory

chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the binuclear center formed by heme A3 and copper B (By similarity).

-!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.

-!- PATHWAY: Respiratory chain; terminal step.

-!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).

-!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family. EMBL; AF466951; AAL75524.1; -.

EMBL; AF466952; AAL75525.1; -.

EMBL; AF466953; AAL75526.1; -.

EMBL; AF466954; AAL75527.1; -.

EMBL; AF466955; AAL75528.1; -.

GO; GO:0019866; C:inner membrane; IEA.

GO; GO:0016021; C:integral to membrane; IEA.

GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

GO; GO:0005739; F:cytochrome-c oxidase activity; IEA.

GO; GO:0016491; F:oxidoreductase activity; IEA.

GO; GO:0006118; P:electron transport; IEA.

GO; GO:0006810; P:transport; IEA.

InterPro; IPR000883; COX1.

Pfam; PF00115; COX1; 1.

PRINTS; PRO1165; CYCOXIDASE1.

PROSITE; PS00077; COX1; 1.

Copper; Electron transport; Heme; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.

NON TER 1 1

FT NON TER 349 349

SQ SEQUENCE 349 AA; 62579F8A877BF54A CRC64;

Query Match 30.0%; Score 6; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
| | | | |
Db 177 SLWALG 182

RESULT 65
Q8SEJ2 PRELIMINARY; PRT; 349 AA.

ID Q8SEJ2; 21, Created)

DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)

DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)

DE Cytochrome oxidase subunit I (Fragment).

GN Name=COI;

OS Palame crassimana.

OG Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia; Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthocnini; Palame.

OC NCBI_TaxID=185373;

OX [1]

RN SEQUENCE FROM N.A.

RP Berkov A.;

RT "The impact of redefined species limits in Palame (Coleoptera, Cerambycidae, Lamiinae, Acanthocnini) on assessments of host, seasonal, and stratum specificity."

RL Biol. J. Linn. Soc. Lond. 76:195-209(2002).

-!- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2

and heme A of subunit 1 to the binuclear center formed by heme A3 and copper B (By similarity).

-!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.

-!- PATHWAY: Respiratory chain; terminal step.

-!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).

-!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family. EMBL; AF466941; AAL75514.1; -.

EMBL; AF466942; AAL75515.1; -.

EMBL; AF466943; AAL75516.1; -.

EMBL; AF466944; AAL75517.1; -.

EMBL; AF466945; AAL75518.1; -.

EMBL; AF466946; AAL75519.1; -.

EMBL; AF466947; AAL75520.1; -.

EMBL; AF466948; AAL75521.1; -.

EMBL; AF466949; AAL75522.1; -.

EMBL; AF466950; AAL75523.1; -.

GO; GO:0019866; C:inner membrane; IEA.

GO; GO:0016021; C:integral to membrane; IEA.

GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

GO; GO:0005739; F:cytochrome-c oxidase activity; IEA.

GO; GO:0016491; F:oxidoreductase activity; IEA.

GO; GO:0006118; P:electron transport; IEA.

GO; GO:0006810; P:transport; IEA.

InterPro; IPR000883; COX1.

Pfam; PF00115; COX1; 1.

PRINTS; PRO1165; CYCOXIDASE1.

PROSITE; PS00077; COX1; 1.

Copper; Electron transport; Heme; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.

NON TER 1 1

FT NON TER 349 349

SQ SEQUENCE 349 AA; 5D199239F87C90E6 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
| | | | |
Db 177 SLWALG 182

RESULT 66
Q8SEJ3 PRELIMINARY; PRT; 349 AA.

ID Q8SEJ3; 21, Created)

DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)

DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)

DE Cytochrome oxidase subunit I (Fragment).

GN Name=COI;

OS Palame crassimana.

OG Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia; Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthocnini; Palame.

OC NCBI_TaxID=185373;

OX [1]

RN SEQUENCE FROM N.A.

RP Berkov A.;

RT "The impact of redefined species limits in Palame (Coleoptera, Cerambycidae, Lamiinae, Acanthocnini) on assessments of host, seasonal, and stratum specificity."

RL Biol. J. Linn. Soc. Lond. 76:195-209(2002).

-!- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in

```

CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the bimetallic center formed by heme A3
CC and copper B (By similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC
CC -1- PATHWAY: Respiratory chain; terminal step.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AF466939; AAL75512.1; -.
DR EMBL; AF466940; AAL75513.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0006118; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
DR Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON_TER 1 349
FT NON_TER 349 349
SQ SEQUENCE 349 AA; 39075 MW; 5D068D39E76390E6 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
DB 177 SLWALG 182
|||||

RESULT 67
Q8SEP8 PRELIMINARY; PRT; 349 AA.
AC Q8SEP8;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Palame mimetica.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthociniini;
OC Palame.
OX NCBI_TaxID=185376;
RN [1]
RP SEQUENCE FROM N.A.
RA Berkov A.;
RT "The impact of redefined species limits in Palame (Coleoptera,
RT Cerambycidae, Lamiinae, Acanthociniini) on assessments of host,
RT seasonal, and stratum specificity."
RL Biol. J. Linn. Soc. Lond. 76:195-209(2002).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme. Electrons originating in
CC catalytic subunit of the enzyme. Electrons originating in
CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the bimetallic center formed by heme A3
CC and copper B (By similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC
CC -1- PATHWAY: Respiratory chain; terminal step.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AF466939; AAL75512.1; -.
DR EMBL; AF466940; AAL75513.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0006118; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
DR Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON_TER 1 349
FT NON_TER 349 349
SQ SEQUENCE 349 AA; 39075 MW; 5D068D39E76390E6 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
DB 177 SLWALG 182
|||||

RESULT 68
Q8SKES PRELIMINARY; PRT; 349 AA.
AC Q8SKES;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Sporetus seminalis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthociniini;
OC Sporetus.
OX NCBI_TaxID=185381;
RN [1]
RP SEQUENCE FROM N.A.
RA Berkov A.;
RT "The impact of redefined species limits in Palame (Coleoptera,
RT Cerambycidae, Lamiinae, Acanthociniini) on assessments of host,
RT seasonal, and stratum specificity."
RL Biol. J. Linn. Soc. Lond. 76:195-209(2002).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. CO I is the
CC catalytic subunit of the enzyme. Electrons originating in
CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the bimetallic center formed by heme A3
CC and copper B (By similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC
CC -1- PATHWAY: Respiratory chain; terminal step.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AF466939; AAL75512.1; -.
DR EMBL; AF466940; AAL75513.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.

```

```

CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AF466977; AAL75550.1; -.
DR EMBL; AF466978; AAL75551.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
DR Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON_TER 1 349
FT NON_TER 349 349
SQ SEQUENCE 349 AA; 38978 MW; 2D3815E9B8B1AECC CRC64;

Query Match 30.0%; Score 6; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
DB 177 SLWALG 182
|||||

RESULT 68
Q8SKES PRELIMINARY; PRT; 349 AA.
AC Q8SKES;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit I (Fragment).
GN Name=COI;
OS Sporetus seminalis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthociniini;
OC Sporetus.
OX NCBI_TaxID=185381;
RN [1]
RP SEQUENCE FROM N.A.
RA Berkov A.;
RT "The impact of redefined species limits in Palame (Coleoptera,
RT Cerambycidae, Lamiinae, Acanthociniini) on assessments of host,
RT seasonal, and stratum specificity."
RL Biol. J. Linn. Soc. Lond. 76:195-209(2002).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. CO I is the
CC catalytic subunit of the enzyme. Electrons originating in
CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the bimetallic center formed by heme A3
CC and copper B (By similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC
CC -1- PATHWAY: Respiratory chain; terminal step.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AF466939; AAL75512.1; -.
DR EMBL; AF466940; AAL75513.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.

```

DR GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006118; P:electron transport; IEA.
 DR GO:0006810; P:transport; IEA.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.
 DR PRINTS: PR01165; CYCOXIDASE1.
 DR PROSITE: PS00077; COX1; 1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 1
 FT NON_TER 349 349
 SQ SEQUENCE 349 AA; 39138 MW; BAB5303A8AB7926E CRC64;

Query Match 30.0%; Score 6; DB 2; Length 349;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 Db 177 SLWALG 182

RESULT 69
 Q8SKE6
 ID Q8SKE6 PRELIMINARY; PRT; 349 AA.
 AC Q8SKE6;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Cytochrome oxidase subunit I (Fragment).
 GN Name=COI;
 OS Oedopeza leucostigma.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthociniini;
 OC Oedopeza.
 OX NCBI_TaxID=185383;
 RN [1]_TaxID=185383;
 RP SEQUENCE FROM N.A.
 RA Berkov A.;
 RT "The impact of redefined species limits in Palame (Coleoptera, Cerambycidae, Lamiinae, Acanthociniini) on assessments of host, seasonal, and stratum specificity";
 RT Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthociniini;
 RL Biol. J. Linn. Soc. Lond. 76:195-209(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the binuclear center formed by heme A3 and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferrocycytochrome c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family. EMBL; AF466980; AAL75553.1; -.
 DR GO:0019866; C:inner membrane; IEA.
 DR GO:0016021; C:integral to membrane; IEA.
 DR GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO:0005739; C:cytochrome c oxidase activity; IEA.
 DR GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006118; P:electron transport; IEA.
 DR GO:0006810; P:transport; IEA.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.
 DR PRINTS: PR01165; CYCOXIDASE1.
 DR PROSITE: PS00077; COX1; 1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 1
 FT NON_TER 349 349
 SQ SEQUENCE 349 AA; 39139 MW; 4341C14E98D6012 CRC64;
 PROSITE; PS00077; COX1; 1.

Query Match 30.0%; Score 6; DB 2; Length 349;

KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 1
 FT NON_TER 349 349
 SQ SEQUENCE 349 AA; 38953 MW; 03815114721FABE9 CRC64;
 Query Match 30.0%; Score 6; DB 2; Length 349;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 Db 177 SLWALG 182

RESULT 70
 Q8SKE7
 ID Q8SKE7 PRELIMINARY; PRT; 349 AA.
 AC Q8SKE7;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Cytochrome oxidase subunit I (Fragment).
 GN Name=COI;
 OS Neoeutrypanus sp. AMNH-1806.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Chrysomeloidea; Cerambycidae; Lamiinae; Acanthociniini;
 OC Neoeutrypanus.
 OX NCBI_TaxID=185386;
 RN [1]_TaxID=185386;
 RP SEQUENCE FROM N.A.
 RA Berkov A.;
 RT "The impact of redefined species limits in Palame (Coleoptera, Cerambycidae, Lamiinae, Acanthociniini) on assessments of host, seasonal, and stratum specificity";
 RT Biol. J. Linn. Soc. Lond. 76:195-209(2002).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the binuclear center formed by heme A3 and copper B (By similarity).
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferrocycytochrome c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family. EMBL; AF466979; AAL75552.1; -.
 DR GO:0019866; C:inner membrane; IEA.
 DR GO:0016021; C:integral to membrane; IEA.
 DR GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO:0005739; C:cytochrome c oxidase activity; IEA.
 DR GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006118; P:electron transport; IEA.
 DR GO:0006810; P:transport; IEA.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.
 DR PRINTS: PR01165; CYCOXIDASE1.
 DR PROSITE: PS00077; COX1; 1.
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
 KW Transport.
 FT NON_TER 1 1
 FT NON_TER 349 349
 SQ SEQUENCE 349 AA; 39139 MW; 4341C14E98D6012 CRC64;
 PROSITE; PS00077; COX1; 1.

Query Match 30.0%; Score 6; DB 2; Length 349;

```
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
Db 177 SLWALG 182

RESULT 71
Q8RDG5
ID Q8RDG5 PRELIMINARY; PRT; 349 AA.
AC Q8RDG5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=TFE0067;
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE012981; AAM23374.1; -.
DR InterPro; IPR008928; Glyco trans 6hp.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 349 AA; 40480 MW; 6556C1858C903F8B CRC64;

Query Match 30.0%; Score 6; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
Db 102 SLWALG 107

RESULT 72
AAR01218
ID AAR01218 PRELIMINARY; PRT; 349 AA.
AC AAR01218;
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 14-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE LaeA.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Peizomycolina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5085;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15075281;
RA Box J.W., Keller N.P.;
RT "LaeA, a Regulator of Secondary Metabolism in Aspergillus spp.";
RL Eukaryotic cell 3:527-535(2004).
DR EMBL; AY422723; AAR01218.1; -.
SQ SEQUENCE 349 AA; 40076 MW; 0462C78DLA7E82A0 CRC64;

Query Match 30.0%; Score 6; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
Db 165 SLWALG 170
```

```
RESULT 73
Q9TAQ8
ID Q9TAQ8 PRELIMINARY; PRT; 352 AA.
AC Q9TAQ8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase I (Fragment).
OS Xyleborus affinis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Scolytidae; Xyleborus.
OX NCBI_TaxID=105120;
RN [1]
RP SEQUENCE FROM N.A.
RA Normark B.B., Jordal B.H., Farrell B.D.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. CO I is the
CC catalytic subunit of the enzyme. Electrons originating in
CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the bimetallic center formed by heme A3
CC and copper B (By similarity).
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AF187138; AAF04072.1; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR008883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
FT NON TER 1 1
FT NON TER 352 352
SQ SEQUENCE 352 AA; 38489 MW; 73B86D4ED8F5249E CRC64;

Query Match 30.0%; Score 6; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
Db 232 SLWALG 237

RESULT 74
Q8SHS2
ID Q8SHS2 PRELIMINARY; PRT; 356 AA.
AC Q8SHS2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit I (fragment).
GN Name=COL;
OS Coccolytes cf. distinctus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
```

OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Phytophaga; Scolytidae; Coccotrypes.
 OX NCBI_TaxID=191010;

[1] SEQUENCE FROM N.A.

RX MEDLINE=22065772; PubMed=12069549;
 RA Jordal B.H., Normark B.B., Farrell B.D., Kirkendall L.R.;

RT "Extraordinary haplotype diversity in haplodiploid inbreeders:
 RT phylogenetics and evolution of the bark beetle genus *Coccotrypes*.";

RL Mol. Phylogenetic. Evol. 23:171-188(2002)
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory

CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
 CC 3 form the functional core of the enzyme complex. CO I is the

CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2

CC and heme A of subunit 1 to the binuclear center formed by heme A3
 CC and copper B (by similarity).

CC -!- CATALYTIC ACTIVITY: 4 ferrocycochrome c + O(2) = 4 ferrocycochrome
 CC c + 2 H(2)O.

CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial

CC inner membrane (by similarity).
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.

DR EMBL; AF444066; AAM12695.1; -.
 DR GO; GO:0019866; C:inner membrane; IEA.

DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.

DR GO; GO:0006118; F:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000883; COX1.
 DR Pfam; PF00115; COX1; 1.

DR PRINTS; PR01165; CYCOXIDASE1.
 DR PROSITE; PS00077; COX1; 1.

KW Copper; Electron transport; Heme; Inner membrane; Membrane;
 KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;

KW Transport.
 FT NON_TER 1 356

FT SEQUENCE 356 AA; 39160 MW; 4221C6D0FA1DEA6 CRC64;
 SQ

Query Match 30.0%; Score 6; DB 2; Length 356;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 4 SLWALG 9

Db 248 SLWALG 253

RESULT 75

Q8P802

ID Q8P802 PRELIMINARY; PRT; 356 AA.

AC Q8P802;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE GumG protein.

GN Name=gumG; OrderedLocusNames=XCC2449;

OS Xanthomonas campestris (pv. campestris).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xanthomonas.

OX NCBI_TaxID=340;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 33913 / NCPPB 528;

RX MEDLINE=2202145; PubMed=12024217;

RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,

RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,

RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,

RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,

RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.P.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;

RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";

RL Nature 417:459-463(2002).
 DR EMBL; AE012356; AAM41726.1; -.

DR GO; GO:0016747; F:transferase activity, transferring groups O. . .; IEA.
 DR InterPro; IPR002656; Acyl_transf_3.

DR Pfam; PF01757; Acyl_transf_3; 1.
 KW Complete proteome.

SQ SEQUENCE 356 AA; 39137 MW; D30C0F7B1F40F3BA CRC64;

Query Match 30.0%; Score 6; DB 2; Length 356;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 9 GWRWLR 14

Db 279 GWRWLR 284

Search completed: October 26, 2004, 07:22:20
 Job time : 95.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2004, 06:46:31 ; Search time 85 Seconds
(without alignments)
84.407 Million cell updates/sec

Title: US-10-066-965A-1

Perfect score: 20

Sequence: 1 QVWSLWALGWRWLRYYGWNM 20

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	2	Aar89346 Cdk2-inte
2	20	100.0	20	2	Aaw32121 Interacti
3	20	100.0	20	3	Aay49340 Cdk2 inte
4	13	65.0	20	5	Aac18008 Intracell
5	12	60.0	17	5	Aac18007 Intracell
6	8	40.0	480	7	Aoc82724 Pseudomon
7	6	30.0	13	8	Adl14700 Phosphose
8	6	30.0	26	7	Adh76207 HSP20 pho
9	6	30.0	29	7	Adh76204 heat shoc
10	6	30.0	52	5	Abp08236 Human ORF
11	6	30.0	112	5	Abp64437 Human ORF
12	6	30.0	138	7	Abc084088 Pseudomon
13	6	30.0	171	7	Abc078072 Pseudomon
14	6	30.0	277	7	Adm26747 Hyperther
15	6	30.0	277	7	Abc068821 Pseudomon
16	6	30.0	305	4	Aau87103 Novel cen
17	6	30.0	305	8	Adi54418 Novel hum
18	6	30.0	335	8	Adm90955 Human pha
19	6	30.0	368	5	Abp95597 Human Gpc
20	6	30.0	379	1	Aap70461 Sequence
21	6	30.0	395	7	Abc061423 Klebsiell
22	6	30.0	396	6	Abul1932 Protein e
23	6	30.0	409	4	Aau36106 Klebsiell
24	6	30.0	426	4	Abg14215 Novel hum
25	6	30.0	439	7	Adh87525 Enterococ

26	6	30.0	445	7	ABO68883
27	6	30.0	478	7	ABO73332
28	6	30.0	486	2	AAV39311
29	6	30.0	486	4	AAW70957
30	6	30.0	486	6	ABP57692
31	6	30.0	493	7	ABO61738
32	6	30.0	502	5	ABB93521
33	6	30.0	502	5	AAO15132
34	6	30.0	532	3	AAO2079
35	6	30.0	532	5	AAO15193
36	6	30.0	532	5	AAU97138
37	6	30.0	532	5	AAE16741
38	6	30.0	548	3	RAY44633
39	6	30.0	548	6	AAE14907
40	6	30.0	548	8	ADO77835
41	6	30.0	548	8	ADO77827
42	6	30.0	548	8	ADP44588
43	6	30.0	548	8	ADP44596
44	6	30.0	610	7	ADD15932
45	6	30.0	610	7	ADD15931
46	6	30.0	610	7	ADD15930
47	6	30.0	755	4	AAAG90181
48	6	30.0	755	7	ADL65505
49	6	30.0	800	7	ABO75889
50	6	30.0	880	7	ADD14150
51	6	30.0	896	4	AAW24250
52	6	30.0	984	4	AAW79887
53	6	30.0	984	4	AAW79930
54	6	30.0	1083	4	AAW93160
55	6	30.0	1089	3	AAW99369
56	6	30.0	1089	4	AAW66118
57	6	30.0	1089	4	AAU29156
58	6	30.0	1089	6	ABU58532
59	6	30.0	1089	6	ABU88080
60	6	30.0	1089	6	ABU84395
61	6	30.0	1089	6	ABR66269
62	6	30.0	1089	6	ABR65659
63	6	30.0	1089	6	ABU99599
64	6	30.0	1089	6	ABU82838
65	6	30.0	1089	6	ABU89959
66	6	30.0	1089	6	ABR68208
67	6	30.0	1089	6	ABU96261
68	6	30.0	1089	6	ABU92692
69	6	30.0	1089	6	ABO08769
70	6	30.0	1089	6	ABO02821
71	6	30.0	1089	6	ABR74975
72	6	30.0	1089	6	ABR94737
73	6	30.0	1089	6	ABU85710
74	6	30.0	1089	6	ABU98870
75	6	30.0	1089	6	ABU98085
76	6	30.0	1089	6	ABU91791
77	6	30.0	1089	6	ABU89484
78	6	30.0	1089	6	ABU86325
79	6	30.0	1089	6	ABU67538
80	6	30.0	1089	6	ABU80566
81	6	30.0	1089	6	ABR99484
82	6	30.0	1089	6	ABR98874
83	6	30.0	1089	6	ABO16397
84	6	30.0	1089	6	ABR92297
85	6	30.0	1089	6	ABO18938
86	6	30.0	1089	6	ABR78359
87	6	30.0	1089	6	ABU85095
88	6	30.0	1089	6	ABO00234
89	6	30.0	1089	6	ABO11566
90	6	30.0	1089	6	ABO02211
91	6	30.0	1089	6	ABU88785
92	6	30.0	1089	6	ABU83480
93	6	30.0	1089	6	ABO06281
94	6	30.0	1089	6	ABR59317
95	6	30.0	1089	6	ABO09379
96	6	30.0	1089	6	ABO19243
97	6	30.0	1089	6	ABO11261
98	6	30.0	1089	6	ABR66879

Abc068883	Pseudomon
Abc073332	Pseudomon
Aay39311	SpnO prot
Aab70957	S. spinos
Abp57692	Saccharop
Abc061738	Klebsiell
Abb93521	Herbicida
Aao15132	A thalian
Aao2079	Transcrip
Aao15193	Rhodococc
Aau97138	Rhodococc
Aae16741	Rhodococc
Aay44633	Human org
Aae14907	Human OCT
Ado77835	Rat synap
Ado77827	Human syn
Adp44588	Human syn
Adp44596	Norway ra
Add15932	Aspergill
Add15931	Aspergill
Add15930	Wild-type
Agg90181	C glutami
Adl65505	C. glutam
Abc075889	Pseudomon
Adi14150	Human src
Aam24250	Human EST
Aab79887	Corynebac
Aab79930	Corynebac
Agg93160	C glutami
Aay99369	Human PRO
Aab66118	Protein O
Aau29156	Human PRO
Abu58532	Human PRO
Abu88080	Novel hum
Abu84395	Human sec
AbR66269	Human sec
AbR65659	Human sec
AbU99599	Human PRO
AbU82838	Human PRO
AbU89959	Novel hum
AbR68208	Human sec
AbU96261	Novel hum
AbU92692	Human sec
Abc08769	Human sec
Abc02821	Human sec
AbR74975	Human sec
AbR94737	Human sec
AbU85710	Human PRO
AbU98870	Novel hum
AbU98085	Novel hum
AbU91791	Novel hum
AbU89484	Human PRO
AbU86325	Human sec
AbU67538	Human sec
AbU80566	Human PRO
AbR99484	Human sec
AbR98874	Human sec
Abc16397	Human sec
AbR92297	Human sec
Abc18938	Human sec
AbR78359	Human sec
AbU85095	Novel hum
Abc00234	Novel hum
Abc11566	Human sec
Abc002211	Human sec
AbU88785	Novel hum
AbU83480	Human sec
Abc06281	Novel hum
AbR59317	Human sec
Abc009379	Human sec
Abc19243	Novel hum
Abc11261	Human sec
AbR66879	Human sec

99 6 30.0 1089 6 ABO16092 Abo16092 Human sec
100 6 30.0 1089 6 ABO13798 Abo13798 Human sec

ALIGNMENTS

RESULT 1

AAW32121
ID AAR89346 standard; peptide; 20 AA.

XX AAR89346;

DT 10-SEP-1996 (first entry)

DE Cdk2-interacting peptide isolated using interaction trap assay.

KW Cdk2; cyclin dependent kinase 2; assay; identifying; isolating;
KW cell cycle; interaction; antagonist; conformationally-constrained;
KW agonist; interaction trap; thioredoxin; LexA; two-hybrid system.

OS Synthetic.

XX WO9602561-A1.

XX 01-FEB-1996.

XX 20-JUL-1995; 95WO-US009307.

XX 20-JUL-1994; 94US-00278082.

XX (GEMO) GEN HOSPITAL CORP.

XX (GEMY) GENETICS INST INC.

PI Brent R, McCoy JM, Jessen TH, Xu C;

XX WPI; 1996-105852/11.

XX Interaction trap systems using conformationally-constrained proteins -
PT useful for detection of protein interactions and for identification and
PT isolation of interacting proteins.

XX Claim 66; Page 60; 73pp; English.

XX AAR89340-R89351 are Cdk-2 interacting peptides identified by an
CC interaction trap assay using conformationally-constrained proteins. The
CC assay comprises providing a host cell (esp. a yeast cell) which contains:
CC (i) a reporter gene operably linked to a DNA-binding protein recognition
CC site, pref. LexA; (ii) a 1st fusion gene expressing Cdk-2 covalently
CC bonded to a LexA binding protein; and (iii) a 2nd fusion gene expressing
CC a conformationally-constrained (pref. with thioredoxin) protein (either
CC expected to interact with Cdk-2 or a random peptide) covalently bonded to
CC a gene-activating moiety, and measuring expression of the reporter gene
CC as a measure of interaction between Cdk2 and the conformationally
CC constrained protein. The same system may be applied to find proteins that
CC interact with any other protein of interest (e.g. Ras) or proteins having
CC agonist or antagonist activity on such interactions

SQ Sequence 20 AA;

Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.4e-14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVWSLWALGWRWLRRYGNM 20

DB 1 QVWSLWALGWRWLRRYGNM 20

RESULT 2

AAW32121
ID AAW32121 standard; peptide; 20 AA.

XX

AC AAW32121;

XX 22-APR-1998 (first entry)

XX Interaction trap system aptamer 7.

XX Protein interaction; interaction trap; fusion protein; mimetic;
KW therapeutic; detection; reporter gene.

OS Synthetic.

XX WO9738127-A1.

XX 16-OCT-1997.

XX 09-APR-1997; 97WO-US005793.

XX 09-APR-1996; 96US-00630052.

XX (GEMO) GEN HOSPITAL CORP.

XX (GEMY) GENETICS INST INC.

PI Brent R, McCoy JM, Jessen TH, Xu CW;

XX WPI; 1997-512733/47.

XX New trap system for detecting protein interactions - comprises a reporter
PT gene linked to a DNA-binding-protein recognition site and fusion proteins
PT to test for interactions.

XX Claim 73; Page 52; 89pp; English.

XX Aptamers AAW32116-W32132 have been isolated from a peptide library and
CC are used in a novel interaction trap method for detecting protein
CC interactions and isolating novel proteins. The method involves a host
CC cell containing a reporter gene operably linked to a DNA-binding-protein
CC (DBP) recognition site, a fusion gene capable of expressing a fusion
CC protein which is able to specifically bind to the DBP recognition site
CC and a second fusion gene which expresses a second fusion protein which is
CC conformationally constrained and bonded to a gene activating moiety.
CC Measuring expression of the reporter gene gives a measure of the
CC interaction between P1 and P2. This method can be used to identify
CC agonists or antagonists for use as therapeutic molecules or for the
CC design of simple molecule mimetics. The method is specifically used to
CC detect an interacting protein in a population of proteins or to identify
CC a candidate interactor. Using conformationally constrained proteins can
CC provide for tertiary structural analysis and can also protect proteins
CC from cellular degradation and/or increase the protein's solubility,
CC and/or otherwise alter the capacity of the candidate interactor to
CC interact

SQ Sequence 20 AA;

Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.4e-14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVWSLWALGWRWLRRYGNM 20

DB 1 QVWSLWALGWRWLRRYGNM 20

RESULT 3

AAW49340

ID AAY49340 standard; peptide; 20 AA.

XX AAY49340;

XX 14-MAR-2000 (first entry)

XX Cdk2 interacting peptide i5-4.

XX Saccharomyces; mammalian; fusion protein; interactor peptide;

KW conformation-constraining protein; DNA binding moiety; Cdk2;
 KW gene activating moiety; protein interaction; gene purification.
 XX Synthetic.

OS
 PN US6004746-A.

XX
 PD 21-DEC-1999.

XX
 PF 20-JUL-1995; 95US-00504538.

XX
 PR 20-JUL-1994; 94US-00278082.

XX
 PA (GEMO) GEN HOSPITAL CORP.

XX
 PA (GEMO) GENETICS INST INC.

XX
 PI Mccoy JM, Jessen TH, Brent R;

XX
 DR WPI; 2000-072059/06.

XX
 PT Population of Saccharomyces and/or mammalian cells comprising recombinant
 PT DNA encoding fusion proteins, useful for detecting protein interactions.

XX
 PS Disclosure; Fig 3B; 24pp; English.

XX
 CC The invention relates to a population of Saccharomyces and/or mammalian
 CC cells comprising recombinant DNA molecules encoding fusion proteins, each
 CC consisting of a candidate interactor peptide, a conformation-constraining
 CC protein and a DNA binding moiety and/or gene activating moiety. The cells
 CC are useful for detecting protein interactions. The cells may also be used
 CC in a method for identifying and purifying genes encoding a wide range of
 CC useful proteins based on their physical interaction with a second
 CC polypeptide

XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 20; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.4e-14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVWSLWALGWRWLRYYGWNM 20

DB 1 QVWSLWALGWRWLRYYGWNM 20

RESULT 4

AAO18008

ID AAO18008 standard; peptide; 20 AA.

XX
 AC AAO18008;

XX
 DT 30-AUG-2002 (first entry)

XX
 DE Intracellular target molecule property modulation method aptamer 10M.

XX
 KW Intracellular target; cellular component; property modulation;
 KW antimicrobial; immunomodulatory; nootropic; neuroprotective; metabolic;
 KW neuroleptic; cytostatic; cardiatic; infection; immunological disorder;
 KW neurological disorder; metabolic disorder; psychiatric disorder;
 KW myopathy; cancer; cardiovascular disorder.

XX
 OS Unidentified.

XX
 PN EP1205191-A1.

XX
 PD 15-MAY-2002.

XX
 PF 13-NOV-2000; 2000EP-00403156.

XX
 PR 13-NOV-2000; 2000EP-00403156.

XX
 PA (CNRS) CENT NAT RECH SCI.

XX
 PA (MASS-) MASSACHUSETTS GEN HOSPITAL.

PA (MOLE-) MOLECULAR SCI INST.

XX
 PI Colas P, Brent R, Cohen BA;

XX
 DR WPI; 2002-418829/45.

XX
 PT Process for specifically modulating the properties of an intracellular
 PT target molecule used for the treatment of various disorders.

XX
 PS Example 1; Fig 1; 33pp; English.

XX
 CC The present invention relates to a process for specifically modulating
 CC the properties of an intracellular target molecule T, and/or of a
 CC cellular component C which interacts directly or indirectly in a cell
 CC with the target. The process involves the introduction into the cell of a
 CC chimeric molecule known as a targeted effector, comprising a recognition
 CC moiety capable of recognising T and an effector moiety. The chimeric
 CC protein or nucleic acid can be used in the preparation of a medicament
 CC for the treatment of microbial infections, immunological disorders,
 CC neurological disorders, metabolic disorders, psychiatric disorders,
 CC myopathies, genetic disorders, cancer, cardiovascular disorders and
 CC dental disorders. The present sequence is a mutant of a known anti-Cdk2
 CC aptamer used in the exemplification of the invention

XX
 SQ Sequence 20 AA;

Query Match 65.0%; Score 13; DB 5; Length 20;

Best Local Similarity 100.0%; Pred. No. 9.7e-07;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WALGWRWLRYYGW 18

DB 6 WALGWRWLRYYGW 18

RESULT 5

AAO18007

ID AAO18007 standard; peptide; 17 AA.

XX
 AC AAO18007;

XX
 DT 30-AUG-2002 (first entry)

XX
 DE Intracellular target molecule property modulation method aptamer 10.

XX
 KW Intracellular target; cellular component; property modulation;
 KW antimicrobial; immunomodulatory; nootropic; neuroprotective; metabolic;
 KW neuroleptic; cytostatic; cardiatic; infection; immunological disorder;
 KW neurological disorder; metabolic disorder; psychiatric disorder;
 KW myopathy; cancer; cardiovascular disorder.

XX
 OS Unidentified.

XX
 PN EP1205191-A1.

XX
 PD 15-MAY-2002.

XX
 PF 13-NOV-2000; 2000EP-00403156.

XX
 PR 13-NOV-2000; 2000EP-00403156.

XX
 PA (CNRS) CENT NAT RECH SCI.

XX
 PA (MASS-) MASSACHUSETTS GEN HOSPITAL.

XX
 PA (MOLE-) MOLECULAR SCI INST.

XX
 PI Colas P, Brent R, Cohen BA;

XX
 DR WPI; 2002-418829/45.

XX
 PT Process for specifically modulating the properties of an intracellular
 PT target molecule used for the treatment of various disorders.

XX
 PS Example 1; Fig 1; 33pp; English.

XX The present invention relates to a process for specifically modulating
 CC the properties of an intracellular target molecule T, and/or of a
 CC cellular component C which interacts directly or indirectly in a cell
 CC with the target. The process involves the introduction into the cell of a
 CC chimeric molecule known as a targeted effector, comprising a recognition
 CC moiety capable of recognising T and an effector moiety. The chimeric
 CC protein or nucleic acid can be used in the preparation of a medicament
 CC for the treatment of microbial infections, immunological disorders,
 CC neurological disorders, metabolic disorders, psychiatric disorders,
 CC myopathies, genetic disorders, cancer, cardiovascular disorders and
 CC dental disorders. The present sequence is a known anti-Cdk2 aptamer used
 CC in the exemplification of the invention
 XX
 XX Sequence 17 AA;

Query Match 60.0%; Score 12; DB 5; Length 17;
 Best Local Similarity 100.0%; Pred. No. 9.6e-06;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVWSLWALGWRW 12
 |||||
 Db 1 QVWSLWALGWRW 12

RESULT 6
 ABO82724
 ID ABO82724 standard; protein; 480 AA.

AC ABO82724;

DT 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #14899.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD16295.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 31470; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences AB067826-

CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html

XX Sequence 480 AA;

Query Match 40.0%; Score 8; DB 7; Length 480;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LGWRWLR 15
 |||||
 Db 255 LGWRWLR 262

RESULT 7
 ADL14700

ID ADL14700 standard; peptide; 13 AA.

XX ADL14700;

XX 20-MAY-2004 (first entry)

XX Phosphoserine peptide analogue derived from HSP20 protein SeqID 299.

XX human; contractile rate; heart muscle; heat shock protein 20; HSP20;
 KW cardiac disorder; bradyarrhythmia; bradycardia; congestive heart failure;
 KW stunned myocardium; pulmonary hypertension; diastolic dysfunction;
 KW hypertension; anaemia; hyperthyroidism; pulmonary stenosis;
 KW protein therapy; antiarrhythmic; cardiant; hypotensive.

XX Homo sapiens.

XX Location/Qualifiers

PH Key 4

FT Modified-site /note= "Phosphorylated serine"

FT WO2004017912-A2.

XX 04-MAR-2004.

XX 22-AUG-2003; 2003WO-US026366.

XX 22-AUG-2002; 2002US-0405199P.

XX 21-FEB-2003; 2003US-0448953P.

XX (UYAR-) UNIV ARIZONA STATE.

XX (MEDI-) MEDICAL COLLEGE GEORGIA RES INST.

XX Brophy C, Komalavillas P, Pipkin WJ, Johnston JA, Creazzo TL;

XX Burch J;

XX WPI; 2004-257215/24.

XX Increasing the contractile rate in heart muscle, useful for treating
 PT e.g., congestive heart failure, heart muscle failure, pulmonary
 PT hypertension or bradyarrhythmia, comprises administering one or more heat
 PT shock protein 20.

XX Example 6; SEQ ID NO 299; 116pp; English.

XX This invention relates to novel methods for increasing the contractile
 CC rate in heart muscle. Specifically, it refers to peptides derived from
 CC the small heat shock protein 20 (HSP20), a protein that is highly and
 CC constitutively expressed in muscle tissue. Furthermore, an increased
 CC phosphorylation of HSP20 is associated with the cyclic nucleotide-
 CC dependent relaxation of vascular smooth muscle, and increasing the
 CC contractile rate in heart muscle must also comprise decreasing the heart
 CC muscle relaxation rate. The present invention describes peptides and
 CC methods to treat cardiac disorders including bradyarrhythmia,
 CC bradycardia, congestive heart failure, stunned myocardium, pulmonary
 CC hypertension and diastolic dysfunction. In addition, it can be useful for

CC treating congestive heart failure in individuals suffering from, for
 CC example, hypertension, anaemia, hyperthyroidism and pulmonary stenosis.
 CC Accordingly, these compositions can be used for protein therapy and
 CC exhibit antiarrhythmic, cardiant and hypotensive activities. This peptide
 CC sequence is a phosphoserine peptide analogue derived from the human HSP20
 CC protein of the invention.

XX
 SQ Sequence 13 AA;

Query Match 30.0%; Score 6; DB 8; Length 13;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||
 Db 4 SLWALG 9

RESULT 8
 ADH76207
 ID ADH76207 standard; peptide; 26 AA.

XX
 AC ADH76207;
 XX
 DT 22-APR-2004 (first entry)

XX HSP20 phosphopeptide analogue PTD-scrambled-pHSP20 SEQ ID NO:308.

XX heat shock protein 20; cytostatic; antiarteriosclerotic; vasotropic;
 KW antianginal; cerebroprotective; antiarrhythmic; antiasthmatic;
 KW gynaecological; hypotensive; antimigraine; tocolytic; relaxant; HSP;
 KW smooth muscle cell; smooth muscle cell.

XX Synthetic.

XX
 FH Key Location/Qualifiers
 FT Modified-site 17 /note= "Phosphorylated"

XX WO2003018758-A2.

XX 06-MAR-2003.

XX 23-AUG-2002; 2002WO-US026918.

XX 23-AUG-2001; 2001US-0314535P.

XX (UYAR-) UNIV ARIZONA.

XX (BROP/) BROPHY C.

XX (KOMA/) KOMALAVILAS P.

XX (PANI/) PANITCH A.

XX (SEAL/) SEAL B.

XX (LOKE/) LOKESH J.

XX Brophy C, Komalavilas P, Panitch A, Seal B, Lokesh J;

XX WPI; 2003-393248/37.

XX New heat shock protein 20-derived polypeptides, useful for inhibiting,

XX treating or preventing smooth muscle cell vasospasm or a disorder such as

XX intimal hyperplasia, restenosis, atherosclerosis or smooth muscle cell

XX tumors.

XX Example 6; SEQ ID NO 308; 194pp; English.

XX The invention relates to a novel polypeptide comprising a heat shock

XX protein 20-derived polypeptide. A polypeptide of the invention has

XX cytostatic, antiarteriosclerotic, vasotropic, antianginal,

XX cerebroprotective, antiarrhythmic, antiasthmatic, gynaecological,

XX hypotensive, antimigraine, tocolytic, and relaxant activity, and may act

XX as a HSP agonist or antagonist. The polypeptides, heat shock protein

CC atherosclerosis, smooth muscle cell tumors such as leiomyosarcoma, or
 CC vasospasm, which is associated with angina, coronary vasospasm,
 CC Prinzmetal's angina, coronary ischaemia, stroke, bradycardia,
 CC hypertension, pulmonary (lung) hypertension, asthma (bronchospasm),
 CC toxemia of pregnancy, pre-term labour, pre-eclampsia/eclampsia, Raynaud's
 CC disease, Raynaud's phenomenon, haemolytic-uremia, non-occlusive
 CC mesenteric ischaemia, anal fissure, achalasia, impotence, migraine, or
 CC ischaemic muscle injury associated with smooth muscle spasm. The
 CC polypeptides and HSP 20 are also useful for inhibiting smooth muscle cell
 CC proliferation and/or migration. The present sequence represents a peptide
 CC used in a polypeptide of the invention.

XX Sequence 26 AA;

Query Match 30.0%; Score 6; DB 7; Length 26;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||
 Db 17 SLWALG 22

RESULT 9
 ADH76204
 ID ADH76204 standard; peptide; 29 AA.

XX
 AC ADH76204;

XX
 DT 22-APR-2004 (first entry)

XX heat shock protein 20 analogue SEQ ID NO:305.

XX heat shock protein 20; cytostatic; antiarteriosclerotic; vasotropic;
 KW antianginal; cerebroprotective; antiarrhythmic; antiasthmatic;
 KW gynaecological; hypotensive; antimigraine; tocolytic; relaxant; HSP;
 KW smooth muscle cell; smooth muscle cell.

XX Synthetic.

XX
 FH Key Location/Qualifiers
 FT Modified-site 20 /note= "Phosphorylated"

XX WO2003018758-A2.

XX 06-MAR-2003.

XX 23-AUG-2002; 2002WO-US026918.

XX 23-AUG-2001; 2001US-0314535P.

XX (UYAR-) UNIV ARIZONA.

XX (BROP/) BROPHY C.

XX (KOMA/) KOMALAVILAS P.

XX (PANI/) PANITCH A.

XX (SEAL/) SEAL B.

XX (LOKE/) LOKESH J.

XX Brophy C, Komalavilas P, Panitch A, Seal B, Lokesh J;

XX WPI; 2003-393248/37.

XX New heat shock protein 20-derived polypeptides, useful for inhibiting,

XX treating or preventing smooth muscle cell vasospasm or a disorder such as

XX intimal hyperplasia, restenosis, atherosclerosis or smooth muscle cell

XX tumors.

XX Example 5; SEQ ID NO 305; 194pp; English.

XX The invention relates to a novel polypeptide comprising a heat shock

XX protein 20-derived polypeptide. A polypeptide of the invention has

XX cytostatic, antiarteriosclerotic, vasotropic, antianginal,

CC cerebroprotective, antiarrhythmic, antiasthmatic, gynaecological,
 CC hypotensive, antimigraine, tocolytic, and relaxant activity, and may act
 CC as a HSP agonist or antagonist. The polypeptides, heat shock protein
 CC (HSP) 20, and methods are useful for treating or preventing a disorder,
 CC e.g. intimal hyperplasia, stenosis, restenosis, transplant vasculopathy,
 CC atherosclerosis, smooth muscle cell tumors such as leiomyosarcoma, or
 CC vasospasm, which is associated with angina, coronary vasospasm,
 CC Prinzmetal's angina, coronary ischaemia, stroke, bradycardia,
 CC hypertension, pulmonary (lung) hypertension, asthma (bronchospasm),
 CC toxemia of pregnancy, pre-term labour, pre-eclampsia/eclampsia, Raynaud's
 CC disease, Raynaud's phenomenon, haemolytic-uremia, non-occlusive
 CC mesenteric ischaemia, anal fissure, achalasia, impotence, migraine, or
 CC ischaemic muscle injury associated with smooth muscle spasm. The
 CC polypeptides and HSP 20 are also useful for inhibiting smooth muscle cell
 CC proliferation and/or migration. The present sequence represents a peptide
 CC used in a polypeptide of the invention.

XX Sequence 29 AA;

Query Match 30.0%; Score 6; DB 7; Length 29;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 Db 20 SLWALG 25
 |||||

RESULT 10

ABP08236
 ID ABP08236 standard; protein; 52 AA.

XX AC ABP08236;

XX DT 24-JUN-2002 (first entry)

XX DE Human ORFX protein sequence SEQ ID NO:16454.

XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

XX OS Homo sapiens.

XX PN WO200192523-A2.

XX PD 06-DEC-2001.

XX PF 29-MAY-2001; 2001WO-US010836.

XX PR 30-MAY-2000; 2000US-0206132P.

XX PR 29-AUG-2000; 2000US-0228716P.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach MD;

XX XX WPI; 2002-106308/14.

DR N-PSDB; ABN23988.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.

XX PS Disclosure; SEQ ID NO 16454; 1037pp; English.

XX The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1

CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 52 AA;

Query Match 30.0%; Score 6; DB 5; Length 52;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVWSLW 6
 Db 31 QVWSLW 36
 |||||

RESULT 11

ABP64437
 ID ABP64437 standard; protein; 112 AA.

XX AC ABP64437;

XX DT 04-NOV-2002 (first entry)

XX DE Human ORF807.

XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
 KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
 KW cancer; cardiovascular disease; allergy; autoimmune disease;
 KW wound healing; blood coagulation disorder; inflammatory disorder.

XX OS Homo sapiens.

XX PN US2002082206-A1.

XX PD 27-JUN-2002.

XX PF 30-MAY-2001; 2001US-00867550.

XX PR 30-MAY-2000; 2000US-0208427P.

XX PA (LEAC/) LEACH M D.

PA (MEHR/) MEHRABAN F.

PA (CONL/) CONLEY P B.

PA (TOPP/) TOPPER J N.

PA (LAWD/) LAW D.

XX PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

XX WPI; 2002-626554/67.

DR N-PSDB; ABQ99000.

XX New polypeptide designated ORFX are present in human atherogenic cells
 PT and are useful to prevent and treat ORFX-associated disorders including

PT cancer, allergy, wound healing or autoimmune, cardiovascular or
 PT inflammatory disease.

PS Claim 10; SEQ ID NO 1614; 78bp; English.

XX The present invention relates to novel human ORFX polypeptides and their
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
 CC were discovered in human atherogenic cells, in particular in platelets
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
 CC many other tissues as well. Atherogenic cells are cells which have the
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
 CC nucleic acids are useful for treating or preventing a pathological
 CC condition associated with an ORFX-associated disorder, e.g. cancer,
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
 CC coagulation disorders or inflammatory disorders. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from the USPTO web site at
 CC seqdata.uspto.gov/sequence.html?docID=20020082206

XX seqdata.uspto.gov/sequence.html?docID=20020082206

XX Sequence 112 AA;
 Query Match 30.0%; Score 6; DB 5; Length 112;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 WLRRYG 17
 |||||
 Db 4 WLRRYG 9

RESULT 12

ABO84088
 ID ABO84088 standard; protein; 138 AA.

XX ABO84088;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #16263.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD17659.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 32834; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,

CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html

XX Sequence 138 AA;

Query Match 30.0%; Score 6; DB 7; Length 138;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 WRWLRR 15
 |||||
 Db 49 WRWLRR 54

RESULT 13

ABO78072

ID ABO78072 standard; protein; 171 AA.

XX ABO78072;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #10247.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD11643.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 26818; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html

```

XX SQ Sequence 171 AA;
    Query Match          30.0%; Score 6; DB 7; Length 171;
    Best Local Similarity 100.0%; Pred. No. 1.1e+02;
    Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLR 15
    |||||
Db 8 WRWLR 13

RESULT 14
ADM26747
ID ADM26747 standard; protein; 277 AA.
XX AC ADM26747;
XX DT 20-MAY-2004 (first entry)
XX DE Hyperthermophile Methanopyrus kandleri protein #1353.
XX KW hyperthermophile; protein stability enhancement;
XX KW protein activity enhancement.
XX OS Methanopyrus kandleri.
XX PN W02003076575-A2.
XX PD 18-SEP-2003.
XX PF 04-MAR-2003; 2003WO-US006664.
XX PR 04-MAR-2002; 2002US-0361742P.
XX PR 14-MAY-2002; 2002US-0380423P.
XX PR 16-SEP-2002; 2002US-0410974P.
XX PA (FIDE-) FIDELITY SYSTEMS INC.
XX PA (MALY/) MALYKH A.
XX PI Slesarev AI, Pavlov A, Pavlova N, Kozyavkin S;
XX WPI; 2003-748383/70.
XX DR N-PSDB; ADM27081.
XX DE New isolated nucleic acids encoding any of about 1700 Methanopyrus
PT kandleri proteins, and the encoded proteins, useful as a medicaments or
PT as diagnostic agents.
XX Claim 31; SEQ ID NO 1353; 1023pp; English.
XX The invention comprises the amino acid sequence of proteins from the
CC hyperthermophile Methanopyrus kandleri, the invention also comprises the
CC complete genome from Methanopyrus kandleri. The Methanopyrus kandleri
CC proteins of the invention are useful for enhancing the stability and/or
CC activity of other proteins. The Methanopyrus kandleri genome is useful in
CC a variety of diagnostic and analytical methods. The present amino acid
CC sequence represents a Methanopyrus kandleri protein of the invention.
XX SQ Sequence 277 AA;
    Query Match          30.0%; Score 6; DB 7; Length 277;
    Best Local Similarity 100.0%; Pred. No. 1.1e+02;
    Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLR 15
    |||||
Db 272 WRWLR 277

RESULT 15
ABO68821
ID ABO68821 standard; protein; 277 AA.
XX AC ABO68821;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #996.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/S8.
XX DR N-PSDB; ABD02392.
XX DE Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 17567; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 277 AA;
    Query Match          30.0%; Score 6; DB 7; Length 277;
    Best Local Similarity 100.0%; Pred. No. 1.6e+02;
    Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLR 15
    |||||
Db 265 WRWLR 270

RESULT 16
AAU87103
ID AAU87103 standard; protein; 305 AA.
XX AC AAU87103;
XX DT 05-JUN-2002 (first entry)
XX DE Novel central nervous system protein #13.
XX KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;

```

KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy.

XX Homo sapiens.

XX WO200155318-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001332.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-01981123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229503P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 13-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 01-NOV-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.

PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.

PR 01-DEC-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250160P.

PR 05-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.

PR 06-DEC-2000; 2000US-0251479P.

PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-581633/65.
 DR N-PSDB; ABK43433.
 XX
 PT New isolated nucleic acid encoding a protein for diagnosing, preventing,
 PT treating or ameliorating medical conditions and used as food additives or
 PT preservatives.
 XX
 XX Claim 9; SEQ ID NO 621; 837pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule (I) encoding a
 CC novel central nervous system protein. (I) and polypeptides (III) encoded
 CC by (I), are used to treat a medical conditions and in diagnosis of a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
 CC ankylosing lateral sclerosis, infections caused by bacteria, viruses
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
 CC adenocarcinomas and irritable bowel syndrome, reproductive system
 CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
 CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
 CC leukaemia, disorders involving neovascularisation e.g. malignancies,
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
 CC acute kidney failure and blood related disorders e.g. myocardial
 CC infection. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemoraxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,
 Query Match 30.0%; Score 6; DB 4; Length 305;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 WLRRYG 17
 Db 216 WLRRYG 221
 RESULT 17
 ADI54418
 ID ADI54418 standard; protein; 305 AA.
 XX
 AC ADI54418;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Novel human protein seq id 621.
 XX
 XX neuroprotective; nootropic; antiparkinsonian; anticonvulsant;
 KW antidiabetic; antirheumatic; antiarthritic; dermatological;
 KW antiinflammatory; immunosuppressive; antithyroid; antianemic; vasotropic;
 KW anti-HIV; hepatotropic; virucide; antibacterial; fungicide;
 KW antiparasitic; muscular; gynaecological; gastrointestinal; respiratory;
 KW cardiovascular; antiarteriosclerotic; antiarrhythmic; cardiac;
 KW nephrotropic; litholytic; cytosatic; gene therapy; neural disorder;
 KW Alzheimer's disease; Parkinson's disease; Huntington's chorea;
 KW anyotrophic lateral sclerosis; multiple sclerosis;
 KW
 KW immune system disorder; diabetes; rheumatoid arthritis;
 KW systemic lupus erythematosus; autoimmune thyroiditis; haemolytic anaemia;
 KW inflammatory disorder; ischaemia-reperfusion injury;
 KW inflammatory bowel disease; Crohn's disease; infectious disease;
 KW HIV infection; hepatitis infection; bacterial infection;
 KW fungal infection; parasitic infection; muscular disorder;
 KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
 KW cardiovascular disorder; atherosclerosis; arrhythmia; myocarditis;
 KW renal disorder; acute glomerulonephritis; pyelonephritis;
 KW renal lithiasis; proliferative disorder; cancerous diseases; human.
 XX Homo sapiens.
 OS
 XX US2004018969-A1.
 XX
 XX 29-JAN-2004.
 XX
 XX 17-JAN-2001; 2001US-00764875.
 PF
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220363P.
 PR 28-JUL-2000; 2000US-0220364P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.


```
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-02556719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2004-122079/12.
XX N-PSDE; ADI53820.
XX
XX New polypeptides and nucleic acid molecules, useful for detecting,
XX preventing, diagnosing, prognosticating, treating or ameliorating medical
XX conditions e.g. neural disorders, reproductive disorders or infectious
XX diseases.
XX
XX Claim 11; SEQ ID NO 621; 413pp; English.
XX
XX The invention describes an isolated polypeptide comprising an amino acid
XX sequence at least 90% identical to: a polypeptide fragment, domain,
XX epitope, or full-length protein of any one of 607 amino acid sequences
XX (I) described in the specification; a polypeptide fragment of (I), or the
XX encoded sequence contained in (II), having biological activity; or a
XX variant, allelic variant, or a species homologue of (I). The polypeptides
XX and nucleic acid molecules are useful for detecting, preventing,
XX diagnosing, prognosticating, treating or ameliorating medical conditions
XX such as neural disorders, e.g. Alzheimer's disease, Parkinson's disease,
XX Huntington's chorea, amyotrophic lateral sclerosis or multiple sclerosis,
XX immune system disorders, e.g. diabetes, rheumatoid arthritis, systemic
Query Match 30.0%; Score 6; DB 8; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 WLRRYG 17
Db 216 WLRRYG 221
RESULT 18
ADM90955
ID ADM90955 standard; protein; 335 AA.
XX
XX ADM90955;
XX
XX 03-JUN-2004 (first entry)
XX
XX Human pharmaceutically useful protein SeqID 348.
XX
XX human; cancer; haematopoesis; thrombosis; anaemia;
XX cardiovascular disorder; ischaemic heart disease;
XX acute myocardial infarction; respiratory disease; asthma; pneumonia;
XX cystic fibrosis; chronic renal failure; glomerulopathy;
XX gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;
XX HIV infection; systemic lupus erythematosus; endocrine system;
XX diabetes mellitus; epilepsy; Alzheimer's disease;
XX amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;
XX fungal; parasitic; viral infection; cytostatic; anticoagulant;
XX thrombolytic; antianaemic; cardiant; vasotropic; antiasthmatic;
XX antiinflammatory; nephrotropic; antiulcer; hepatotropic;
XX immunosuppressive; antiallergic; dermatological; antirheumatic;
XX antiarthritic; antidiabetic; anticonvulsant; neuroprotective; nootropic;
XX antipsoriatic; antibacterial; fungicide; antiparasitic; virucidal;
XX gene therapy; vaccine.
XX
```

```
OS Homo sapiens.
XX WO2004020595-A2.
XX 11-MAR-2004.
XX 28-AUG-2003; 2003WO-US027107.
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410951P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411111P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.
PA (DNAF-) DNAFORM KK.
XX Williams LT, Chu K, Lee E, Hestir K;
XX WPI; 2004-257410/24.
XX N-PSDB; ADM90746, ADM91164.
XX New human polynucleotides and polypeptides, useful for diagnosing,
PT preventing and treating proliferative disorders, immune disorders,
PT cardiovascular disorders, or bacterial, fungal, parasitic and viral
PT diseases.
XX Claim 1; SEQ ID NO 348; 254pp; English.
XX This invention relates to novel isolated human polynucleotides and the
CC encoded proteins thereof. Specifically, it refers to proteases, kinases,
CC phosphatases, secreted and transmembrane proteins, as well as the derived
CC peptide fragments, which can be used to develop antibodies and screen for
CC small molecule agonists and antagonists that can modulate their
CC activities. The present invention describes polypeptides,
CC polynucleotides, vectors and host cells useful for diagnosing, preventing
CC and treating proliferative disorders, e.g. cancer, disorders of
CC haemopoiesis such as thrombosis and anaemia, cardiovascular disorders,
CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory
CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the
CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,
CC gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune
CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders
CC of the endocrine system, e.g. diabetes mellitus, central nervous system
CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral
CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,
CC parasitic and viral diseases. Accordingly, they exhibit many various
CC activities including cytostatic, anticoagulant, thrombolytic,
CC antanaemic, cardiant, vasotropic, antiaesthatic, antiinflammatory,
CC nephrotropic, antiulcer, hepatotropic, immunosuppressive, antiallergic,
CC dermatological, antirheumatic, antiarthritic, antidiabetic,
CC anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,
CC fungicide, antiparasitic and virucidal, such that these polynucleotides
CC can be used for gene therapy purposes and the development of appropriate
CC vaccines. This polypeptide is a human protein of the invention.
XX

SQ Sequence 335 AA;
Query Match 30.0%; Score 6; DB 8; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
Db 194 SLWALG 199
|||||

RESULT 19
ABP95597
ID ABP95597 standard; protein; 368 AA.
XX AC ABP95597;
XX DT 06-MAR-2003 (first entry)
XX DE Human GPCR polypeptide SEQ ID NO 4.
XX KW Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
XX drug development; gustatory; taste; fragrance; receptor.
XX OS Homo sapiens.
XX PN WO200216548-A2.
XX PD 28-FEB-2002.
XX PF 30-JUL-2001; 2001WO-IB001446.
XX PR 04-AUG-2000; 2000JP-00237818.
XX PR 13-FEB-2001; 2001JP-00034434.
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX Haga T, Takeda S, Mitaku S;
XX WPI; 2002-304118/34.
XX N-PSDB; ABZ42871.
XX Database global search for G protein-coupled receptors, proteins and
PT encoded genes for studying in vivo signal transduction mechanism and
PT identifying targets for drug development.
XX Claim 10; SEQ ID NO 4; 97pp + Sequence Listing; Japanese.
XX The invention relates to a method for screening G protein-coupled
CC receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-
CC ABP95942) by extracting open-reading frames containing 6-8 transmembrane
CC domains with 250-1000 amino acid residues to give a gene homologous with
CC a known GPCR gene. The receptor proteins and encoded genes are useful for
CC studying in vivo signal transduction mechanism and identifying targets
CC for drug development e.g. based on olfactory and gustatory receptors in
CC form of agonists and antagonists by screening intrinsic and extrinsic
CC ligands as bitter taste inhibitors, taste enhancers and fragrance
CC improvers. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at fp.wipo.int/pub/published_pot_sequences
XX Sequence 368 AA;
Query Match 30.0%; Score 6; DB 5; Length 368;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 WLRRYG 17
Db 86 WLRRYG 91
|||||

RESULT 20
```

```

AAP70461
ID AAP70461 standard; protein; 379 AA.
XX
AC
XX AAP70461;
XX
DT 25-MAR-2003 (revised)
DT 13-FEB-1991 (first entry)
XX
DE Sequence of gpG encoded by segment of Xanthomonas campestris DNA that
DE contains a gene cluster that directs Xanthan biosynthesis.
XX
XX Thickening agent; oil recovery; drilling fluid.
XX
XX Xanthomonas campestris.
OS
XX
PN WO8705938-A.
XX
PD 08-OCT-1987.
XX
PF 24-MAR-1987; 87WO-US000604.
XX
PR 24-MAR-1986; 86US-00842944.
PR 23-MAR-1987; 87US-00029530.
XX
PA (GETT-) GETTY SCI DEV CO.
XX
PI Capage MA, Doherty DH, Betlach MR, Vanderslic RW;
XX
DR WPI; 1987-291651/41.
XX
PT Recombinant DNA prodn. of xanthan gum or its variants - by transforming
PT host cells with vector contg. DNA coding for enzymes involved in
PT polysaccharide synthesis.
XX
PS Example; Fig 12; 149pp; English.
XX
CC Virtually all of the segment of Xanthomonas campestris DNA that contains
CC a gene cluster that directs Xanthan biosynthesis (AAN70753), codes for
CC protein products. Each gene is designated by a letter (see Fig 11) and
CC its protein product is designated by that letter preceded by 'gp'
CC (AAP70455-67). (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 379 AA;
Query Match 30.0%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 9 GWRWLR 14
Db 279 GWRWLR 284
|||||
RESULT 21
ABO61423
ID ABO61423 standard; protein; 395 AA.
XX
AC ABO61423;
XX
XX 29-JUL-2004 (first entry)
XX
XX Klebsiella pneumoniae polypeptide seqid 7940.
DE
XX
XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; vaccine.
XX
XX Klebsiella pneumoniae.
OS
XX
PN US6610836-B1.
XX
XX 26-AUG-2003.
PD
XX
XX 27-JAN-2000; 2000US-00489039.
PF
29-JAN-1999; 99US-0117747P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL, Osborne M;
PI
XX WPI; 2003-895346/82.
DR
XX N-PSDB; ACH94974.
DR
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
PT
XX
XX Disclosure; SEQ ID NO 7940; 932pp; English.
PS
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This is the amino acid sequence of a
XX Klebsiella pneumoniae polypeptide of the invention
XX
SQ Sequence 395 AA;
Query Match 30.0%; Score 6; DB 7; Length 395;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 WLRRYG 17
Db 388 WLRRYG 393
|||||
RESULT 22
ABU19932
ID ABU19932 standard; protein; 396 AA.
XX
AC ABU19932;
XX
XX 19-JUN-2003 (first entry)
DT
XX
XX Protein encoded by Prokaryotic essential gene #5459.
DE
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Borrelia cepacia.
OS
XX
XX WO200277183-A2.
PN
XX
XX 03-OCT-2002.
PD
XX
XX 21-MAR-2002; 2002WO-US009107.
PF
XX
XX 21-MAR-2001; 2001US-00815242.
PR
XX 06-SEP-2001; 2001US-00948993.
PR
XX 25-OCT-2001; 2001US-0342923P.
PR
XX 08-FEB-2002; 2002US-00072851.
PR
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR
XX N-PSDB; ACA23802.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 47856; 1766pp; English.
PS

```

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 396 AA;

Query Match 30.0%; Score 6; DB 6; Length 396;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ALGWRW 12
 |||||
 DB 75 ALGWRW 80

RESULT 24
 AAU36106
 ID AAU36106 standard; protein; 409 AA.
 XX
 AC AAU36106;

XX 14-FEB-2002 (first entry)
 DT
 DE Klebsiella pneumoniae cellular proliferation protein #94.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.

XX Klebsiella pneumoniae.
 XX WO200170955-A2.
 XX
 XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.
 XX
 XX 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.
 DR N-PSDB; AASS3965.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 11699; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
 CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 409 AA;

Query Match 30.0%; Score 6; DB 4; Length 409;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||
 DB 356 SLWALG 361

RESULT 24
 ABG14215
 ID ABG14215 standard; protein; 426 AA.

XX
 AC ABG14215;

XX 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #14206.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS78402.

CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-
CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX
SQ Sequence 445 AA;

Query Match 30.0%; Score 6; DB 7; Length 445;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
Db 125 SLWALG 130
|||||

RESULT 27
ABO73332
ID ABO73332 standard; protein; 478 AA.
XX AC ABO73332;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #5507.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX DR WPI; 2003-615309/58.
XX DR N-PSDB; ABD06903.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 22078; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a *P. aeruginosa* nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-*P. aeruginosa* drugs, as templates for recombinant
XX production of *P. aeruginosa*-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of *P. aeruginosa*-caused
XX infection, and in detection of *P. aeruginosa* sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABO67826-
XX ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX
SQ Sequence 478 AA;

Query Match 30.0%; Score 6; DB 7; Length 478;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 WRWLRR 15
Db 443 WRWLRR 448
|||||

RESULT 28
AAY39311
ID AAY39311 standard; protein; 486 AA.

XX AC AAY39311;
XX DT 01-DEC-1999 (first entry)
XX DE SpnO protein involved in spinosyn biosynthesis.

XX KW Spinosyn biosynthetic enzyme; open reading frame; ORF;
XX KW insecticidal microtides; arachnid; nematode; insect; polyketide;
XX KW polyketide synthase; PKS; extender module; initiator module;
XX KW acyl transferase domain; AT; acyl carrier protein; ACP;
XX KW beta-ketosynthase domain; KS; KR; dehydratase domain; DH;
XX KW enoyl reductase domain; ER; beta-ketoreductase; insecticide;
XX KW dideoxysugar synthesis.
XX OS Saccharopolyspora spinosa.

XX PN WO9946387-A1.

XX PD 16-SEP-1999.

XX PF 16-FEB-1999; 99WO-US003212.

XX PR 09-MAR-1998; 98US-00036987.

XX PA (DOWC) DOW AGROSCIENCES LLC.

XX PI Baltz RH, Broughton MC, Crawford KP, Madduri K, Merlo DJ;
XX PI Treadway PJ, Turner JR, Waldron C;

XX DR WPI; 1999-551414/46.

XX DR N-PSDB; AAZ21501.

XX PT New spinosyn biosynthetic genes from Saccharopolyspora spinosa, useful
XX PT for production of insecticidal spinosyn compounds.

XX PS Claim 1; Page 159-161; 190pp; English.

XX This is the amino acid sequence of the product of the spnO gene. The
XX protein is involved in spinosyn biosynthesis. The spnO gene is one of 23
XX genes and open reading frames contained in an 80kb DNA sequence AAZ21501.
XX Spinosyns are insecticidal microtides which are useful for the control of
XX arachnids, nematodes and insects. Biosynthesis of spinosyns occurs via
XX stepwise condensation and modification of carboxylic acid precursors
XX generating a linear polyketide which is modified further. The DNA
XX sequence contains a central region of approximately 55kb which has
XX homology to the DNA encoding the polyketide synthases (PKS) of known
XX macrolide producers. The spinosyn PKS DNA region consists of 5 ORFs with
XX stop codons at the end of acyl carrier protein (ACP) domains. Together
XX the PKS polypeptides (AAY39297-Y39301), form a complex consisting of an
XX initiator module, spnA, and several extender modules spnB-spnE. The
XX products of the genes present in the region upstream of the PKS genes
XX have been assigned names spnF-spnS AAY39302-Y39315 and are responsible
XX for different modifications in spinosyn biosynthesis. There are also two
XX ORFs ORF15 and ORF16 present immediately upstream of spnS, producing
XX polypeptides AAY39316-Y39317, and two ORFs ORF1 and ORF2 present
XX downstream of the PKS region producing polypeptides AAY39318-Y39319. It
XX is suggested that SpnO is involved in dideoxysugar synthesis during
XX spinosyn biosynthesis. The genes are useful to improve yields of
XX spinosyns, and for creating new spinosyns e.g. by mutagenesis, or

CC interruption of steps in spinosyn biosynthesis. The modified spinosyns
 CC may be a new insect control agent or serve as substrates for further
 CC chemical modification and the creation of new semi-synthetic spinosyns.
 CC The genes are also useful to isolate similar sequences from *S. spinosa* or
 CC other species by hybridization
 XX
 SQ Sequence 486 AA;

Query Match 30.0%; Score 6; DB 2; Length 486;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||
 Db 480 SLWALG 485

RESULT 29
 AAB70957
 ID AAB70957 standard; protein; 486 AA.
 XX
 AC AAB70957;
 XX
 XX 28-AUG-2001 (first entry)
 DT
 DE
 DE S. spinosa protein fragment encoded by ORF10, SEQ ID 26.
 XX
 XX Forosamine; trimethylrhamnose; polyketide synthase; biosynthesis;
 KW spinosyn; polyketide aglycone; transgenic plant; insect resistance;
 KW macrolide; insecticidal; 2,3-dehydratase.
 XX
 XX Saccharopolyspora spinosa.
 OS
 XX
 XX DEL19957268-A1.
 PN
 XX
 XX 08-MAR-2001.
 PD
 XX
 XX 29-NOV-1999; 99DE-01057268.
 PF
 XX
 XX 27-AUG-1999; 99DE-01040596.
 PR
 XX
 XX (FARB) BAYER AG.
 PA
 XX
 XX Eberz G, Moehrlie V, Froede R, Velten R, Salas JA;
 PI
 XX
 XX WPI; 2001-267102/28.
 DR
 DR N-PSDB; AAF88327.
 XX
 XX New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
 PT recombinant production of insecticidal spinosyns and their derivatives.
 PT
 XX
 XX Claim 47; Page 139-140; 354pp; German.
 PS
 XX

CC This invention describes a novel method nucleic acid (I) and its encoded
 CC polypeptide (II) containing at least one region that encodes an enzymatic
 CC activity involved in biosynthesis of spinosyns. (I) are used (i) to
 CC identify, inactivate or modulate genes involved in the biosynthesis of
 CC (II); (ii) to generate a library of polyketide synthases; (iii) for
 CC adding forosamine or trimethylrhamnose to a spinosyn or polyketide
 CC aglycone; and (iv) for recombinant production of the corresponding
 CC enzymes, which are used for production of (II), their precursors or
 CC derivatives, including production of transgenic plants that express (II)
 CC and thus have increased resistance to insects. (I) are also useful as
 CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
 CC macrolides with insecticidal, but not antibacterial, activity, and can
 CC also be used to raise specific antibodies, useful for identifying
 CC expression clones in a gene bank. Cells transformed with (I) may produce
 CC (II) at significantly increased levels or produce new derivatives of
 CC (II). This sequence represents an *S. spinosa* 2,3-dehydratase
 XX

SQ Sequence 486 AA;

Query Match 30.0%; Score 6; DB 4; Length 486;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWALG 9
 |||||
 Db 480 SLWALG 485

RESULT 30
 ABP57692
 ID ABP57692 standard; protein; 486 AA.
 XX
 AC ABP57692;
 XX
 XX 22-JAN-2003 (first entry)
 DT
 DE
 DE Saccharopolyspora busO butenyl-spinosyn biosynthetic gene product.
 XX
 XX Butenyl; biosynthetic enzyme; PKS; polyketide synthetase; macrolide;
 KW metabolite; spinosyn.
 XX
 XX Saccharopolyspora sp.
 OS
 XX
 XX WO200279477-A2.
 PN
 XX
 XX 10-OCT-2002.
 PD
 XX
 XX 28-MAR-2002; 2002WO-US009968.
 PF
 XX
 XX 30-MAR-2001; 2001US-0280175P.
 PR
 XX
 XX (DOWC) DOW AGROSCIENCES LLC.
 PA
 XX
 XX Hahn DR, Jackson JD, Bullard BS, Gustafson GD, Waldron C;
 PI Mitchell JC;
 PI
 XX
 XX WPI; 2003-058434/05.
 DR
 DR N-PSDB; ABV75558.
 XX
 XX New butenyl-spinosyn biosynthetic genes, useful for increasing the
 PT production of butenyl-spinosyn insecticidal macrolides, or for changing
 PT the metabolites or products produced by spinosyn-producing
 PT microorganisms.
 PT
 XX
 XX Claim 1; Page 197-199; 218pp; English.
 PS
 XX
 XX The invention relates to a novel DNA molecule comprising a DNA sequence
 CC that encodes a butenyl-spinosyn biosynthetic enzyme, a butenyl-spinosyn
 CC PKS (polyketide synthetase) domain, or a spinosyn PKS module. The butenyl
 CC -spinosyn biosynthetic genes are useful for increasing the production of
 CC butenyl-spinosyn insecticidal macrolides. The genes are also useful for
 CC changing the metabolites or products produced by spinosyn-producing
 CC microorganisms. The present sequence represents one of the butenyl-
 CC spinosyn biosynthetic polypeptides of the invention
 XX
 XX Sequence 486 AA;
 SQ
 Query Match 30.0%; Score 6; DB 6; Length 486;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SLWALG 9
 |||||
 Db 480 SLWALG 485

RESULT 31
 ABO61738
 ID ABO61738 standard; protein; 493 AA.
 XX
 AC ABO61738;
 XX
 XX 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polypeptide seqid 8255.
DE Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX Klebsiella pneumoniae.
OS US6610836-B1.
PN 26-AUG-2003.
XX 27-JAN-2000; 2000US-00489039.
PF 29-JAN-1999; 99US-0117747P.
XX (GENO-) GENOME THERAPEUTICS CORP.
PA Breton GL, Osborne M;
XX WPI; 2003-895346/82.
DR N-PSDB; ACH95289.
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX Disclosure; SEQ ID NO 8255; 932pp; English.
PS The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
XX Klebsiella pneumoniae polypeptide of the invention
CC
XX Sequence 493 AA;
Query Match 30.0%; Score 6; DB 7; Length 493;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 SLWALG 9
Db 440 SLWALG 445
RESULT 32
ID ABB93521 standard; protein; 502 AA.
XX ABB93521;
AC 31-MAY-2002 (first entry)
XX Herbicidally active polypeptide SEQ ID NO 2732.
DE Herbicidally active polypeptide SEQ ID NO 2732.
XX Herbicidal; plant; agriculture; herbicide.
XX Arabidopsis thaliana.
OS WO200210210-A2.
PN 07-FEB-2002.
XX 28-AUG-2001; 2001WO-EF009892.
XX 28-AUG-2001; 2001WO-EF009892.
XX (FARB) BAYER AG.
XX Tietjen K, Weidler M;
XX

DR WPI; 2002-269010/31.
XX Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX Claim 5; SEQ ID NO 2732; 261pp + Sequence Listing; English.
XX The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX Sequence 502 AA;
XX
XX Query Match 30.0%; Score 6; DB 5; Length 502;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 LRRYGW 18
Db 395 LRRYGW 400
RESULT 33
AAO15132
ID AAO15132 standard; protein; 502 AA.
XX AAO15132;
AC 25-SEP-2002 (first entry)
XX A thaliana GAD1.
XX GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW plant stress; GABA; gamma-aminobutyric acid; stress resistance.
XX Arabidopsis thaliana.
OS WO200238736-A2.
PN 16-MAY-2002.
XX 07-NOV-2001; 2001WO-US047447.
PF 07-NOV-2000; 2000US-0246367P.
XX (EMER-) EMERALD BIOAGRICULTURE CORP.
PA Kinnersley AM, Turano FU;
XX WPI; 2002-490073/52.
DR N-PSDB; AAL43410.
XX Making transformed plants that selectively increase gamma-aminobutyric
PT acid production, by incorporating a DNA construct with a polynucleotide
PT encoding a plant glutamic acid decarboxylase enzyme into plant's genome.
XX Claim 17; Page 53; 63pp; English.
XX The present invention relates to a method of producing a transformed
CC plant that selectively increases production of gamma-aminobutyric acid
CC (GABA) in response to a signal, by incorporating into the plant's genome
CC a DNA construct with a non-constitutive promoter operably linked to a
CC polynucleotide encoding a functional plant glutamic acid decarboxylase
CC (GAD), to provide a transformed plant that expresses the GAD coding
CC sequence in response to a signal. Plants of this type have an enhanced

CC ability to tolerate environmental or other stresses. The present sequence
 CC is the A. thaliana GAD1 protein
 XX
 SQ Sequence 502 AA;

Query Match 30.0%; Score 6; DB 5; Length 502;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRYCW 18
 |||||
 Db 395 LRRYCW 400

RESULT 34
 AAB02079
 ID AAB02079 standard; protein; 532 AA.

XX
 AC AAB02079;

XX
 DT 12-SEP-2003 (revised)
 DT 03-JAN-2001 (first entry)

DE Transcription factor partial sequence.

XX Identification; prokaryote; polymerase chain reaction; PCR;
 KW amplification; primer; differential display; picric acid degradation;
 KW gene cluster; open reading frame; ORF; dehydratase; dehydrogenase;
 KW transcription factor; Acyl-CoA synthase; NADPH oxidoreductase.

XX Rhodococcus erythropolis; HL PM-1.

XX
 FH Key Location/Qualifiers
 FT Misc-difference 532
 FT /note= "Gln or Glu"

XX WO2000049177-A2.

XX 24-AUG-2000.

XX 17-FEB-2000; 2000WO-US003989.

XX 19-FEB-1999; 99US-0120702P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Rouviere P;

XX WPI; 2000-587069/55.

XX N-PSDB; AAA53942.

XX Differential display method using a large number of arbitrary primers for
 RT-PCR used to isolate novel differentially expressed prokaryotic genes.

XX Disclosure; Page 46-47; 66pp; English.

XX A new method to identify differentially expressed prokaryotic genes using
 CC a large number of arbitrarily primed polymerase chain reactions comprises
 CC separating two populations of microbial cells, where a first population
 CC is contacted with a stimulating agent; extracting total RNA from both
 CC microbial cell populations; amplifying the extracted RNA from both
 CC populations by preparing a collection of at least thirty-two different
 CC arbitrary primers, where each primer comprises a common and a variable
 CC region; individually contacting each primer of with a sample of extracted
 CC RNA from the two populations under conditions where two sets of
 CC amplification products are produced; purifying the two sets of
 CC amplification products; identifying the amplification products generated
 CC in the first population which differ from products generated from the
 CC second population as differentially expressed genes; and optionally
 CC sequencing the identified differentially expressed genes. The advantage
 CC over previous methods is that previous methods of differential display to
 CC clone genes using thirty-two or thirty primers have isolated four and one
 CC genes, respectively. The new method using a greater number of primers has

CC isolated twenty-one induced gene fragments. This sequence is a partial
 CC sequence of the transcription factor encoded by open reading frame (ORF)
 CC 1 of the picric acid degradation cluster (See GENESEQ record AAA53941).
 CC (Updated on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 532 AA;

Query Match 30.0%; Score 6; DB 3; Length 532;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||
 Db 332 SLWALG 337

RESULT 35
 AAO15193
 ID AAO15193 standard; protein; 532 AA.

XX
 AC AAO15193;

XX 05-SEP-2002 (first entry)

XX Rhodococcus picric acid degradation pathway-related transcription factor.

XX Picric acid degradation gene cluster; recombinant organism;

XX picric acid degradation pathway; transcription factor.

XX Rhodococcus erythropolis.

XX Key Location/Qualifiers

XX Misc-difference 532

XX /note= "Encoded by TGA"

XX US2002042117-A1.

XX 11-APR-2002.

XX 17-SEP-2001; 2001US-00955597.

XX 03-SEP-1999; 99US-0152545P.

XX 31-AUG-2000; 2000US-00651941.

XX (ROUVI) ROUVIERE P E.

XX (WALT/) WALTERS D M.

XX (RUSS/) RUSS R.

XX Rouviere PE, Walters DM, Russ R;

XX WPI; 2002-381946/41.

XX N-PSDB; AAL43620.

XX Isolated nucleic acid fragments encoding enzymes of the picric acid
 degradation pathway isolated from Rhodococcus erythropolis HL PM-1,
 XX useful in the creation of recombinant organisms that have the ability to
 XX degrade picric acid.

XX Disclosure; Page 23-24; 53pp; English.

XX The invention comprises 12 Rhodococcus erythropolis ORFs encoding enzymes
 CC of the picric acid degradation pathway. The invention also comprises the
 CC nucleotide sequence of the picric acid degradation gene cluster
 CC containing all 12 of the ORFs. The picric acid degradation pathway genes
 CC and enzymes of the invention are useful for creating recombinant
 CC organisms that have the ability to degrade picric acid. As well as for
 CC the identification of new species of bacteria that have the ability to
 CC degrade picric acid. The present amino acid sequence represents a
 CC Rhodococcus erythropolis picric acid degradation pathway-related
 CC transcription factor

XX Sequence 532 AA;

Query Match 30.0%; Score 6; DB 5; Length 532;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 Db 332 SLWALG 337
 |||||

RESULT 36
 AAU97138
 ID AAU97138 standard; protein; 532 AA.
 XX
 AC AAU97138;
 XX
 DT 29-AUG-2003 (revised)
 DT 13-AUG-2002 (first entry)
 XX
 DE Rhodococcus erythropolis HL PM-1 transcription factor.
 XX
 KW Picric acid degradation; 2,4,6-trinitrophenol; explosive manufacturing;
 KW aniline; colour fast dye; pharmaceutical; steel etching;
 KW environmental toxicant; enzymatic degradative process;
 KW transcription factor.
 XX
 OS Rhodococcus erythropolis; strain HL-PM1.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 531..532
 FT /note= "Encoded by GAC"
 XX
 XX US6355470-B1.
 XX
 PD 12-MAR-2002.
 XX
 XX 31-AUG-2000; 2000US-00651941.
 XX
 PF 03-SEP-1999; 99US-0152545P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Rouviere PE, Walters DM, Russ R;
 XX
 XX WPI; 2002-433274/46.
 XX
 XX N-PSDB; ABK51838.

Nucleic acid encoding an F420/NADPH oxidoreductase isolated from
 Rhodococcus erythropolis HL PM-1 is associated with picric acid
 degradation and is useful to create recombinant organisms that degrade.
 Example 6; Col 41-46; 49pp; English.

The present invention relates to the isolation of Rhodococcus
 erythropolis HL PM-1 gene cluster containing 12 open reading frames
 (ORFs) implicated in the degradation of picric acid (2,4,6-
 trinitrophenol). The polynucleotide sequences of the invention are useful
 for creating recombinant organisms that have the ability to degrade
 picric acid. Picric acid which is used in industrial applications
 including the manufacture of explosives, aniline, colour fast dyes,
 pharmaceuticals and in steel etching, is highly unstable. The present
 invention provides a means of disposal/removal of this toxic substance
 from the environment by an enzymatic degradative process. The present
 sequence represents a transcription factor encoded by picric acid
 degradation gene cluster ORF1. (Updated on 29-AUG-2003 to standardise OS
 field)

SQ Sequence 532 AA;

Query Match 30.0%; Score 6; DB 5; Length 532;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||

Db 332 SLWALG 337
 |||||

RESULT 37
 AAE16741
 ID AAE16741 standard; protein; 532 AA.
 XX
 AC AAE16741;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Rhodococcus erythropolis HL-PM1 transcription factor.
 XX
 KW Picric acid degradation gene; cyclohexanol degradation; heavy metal;
 KW inhibitory effect; chemical; environmental pollutant; anaerobiosis;
 KW oxidative damage; pathogenesis; transcription factor.
 XX
 OS Rhodococcus erythropolis.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 531..532
 FT /note= "Encoded by GAC"
 FT Misc-difference 532
 FT /label= Glu, Gln
 XX
 XX US6329151-B1.
 XX
 PD 11-DEC-2001.
 XX
 PF 05-SEP-2000; 2000US-00655270.
 XX
 PR 03-SEP-1999; 99US-0152542P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Rouviere PE;
 XX
 XX WPI; 2002-121127/16.
 XX
 XX N-PSDB; AAD27241.

Identifying differentially expressed genes, by amplifying total RNA of
 first microbial cell population that is contacted with stimulating agent
 and of a second population using arbitrary primers, and comparing them.
 Example 6; Col 39-42; 51pp; English.

The invention relates to a reliable and rapid method to identify
 differentially expressed genes in microbes. The method relies on the use
 of a large number of arbitrarily primed PCR reactions. The method is
 useful for identifying differentially expressed genes in microbes, and
 for distinguishing genetic differences between two populations of cells
 which differ in genotype. This method is useful for identifying the DNA
 sequences of genes involved in the degradation of the picric acid from
 Rhodococcus erythropolis strain HL PM-1, and genes involved in
 cyclohexanol degradation from a consortium of organisms, or to detect
 cDNA fragments from differentially expressed mRNAs. This method is useful
 for examining the inhibitory effects of various treatments such as
 chemicals, environmental pollutants, heavy metals, changes in
 temperature, changes in pH, agents producing oxidative damage, agents
 producing DNA damage, anaerobiosis, pathogenesis, and changes in nitrate
 availability on mRNA levels. The present sequence is a transcription
 factor encoded by Rhodococcus erythropolis strain HL PM-1 picric acid
 degradation gene ORF1

SQ Sequence 532 AA;

Query Match 30.0%; Score 6; DB 5; Length 532;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLWALG 9
 |||||

```

Db      332 SLWALG 337
RESULT 38
AAV44633
ID      AAV44633 standard; protein; 548 AA.
XX
AC      AAV44633;
XX
XX
DT      07-APR-2000 (first entry)
XX
DE      Human organic cation transporter-like protein (OCTlp).
XX
KW      Human; organic cation transporter-like protein; OCTlp; transporter;
KW      transmembrane; nootropic; neuroprotective; neuroleptic; anticonvulsant;
KW      antiParkinsonian; antidepressant; cellular process; cell proliferation;
KW      screen; treatment; prevention; diagnosis; neurodegenerative disorder;
KW      Alzheimer; Parkinson's; Huntington; ALS; amyotrophic lateral sclerosis;
KW      CNS disorder; central nervous system; schizophrenia; depression;
KW      behavioural; sleep disorder; Alzheimer's; eating disorder.
XX
OS      Homo sapiens.
XX
XX
FH      Key
FH      Location/Qualifiers
FT      1..85
FT      /label= Cytoplasmic_domain
FT      Region
FT      71..524
FT      /note= "shows homology to a consensus sequence for sugar
FT      and other transporter molecules derived from a hidden
FT      Markov model"
FT      86..108
FT      /label= Transmembrane_domain
FT      109..121
FT      /label= Extracellular_domain
FT      /note= "Hydrophilic region useful for antibody
FT      production"
FT      122..140
FT      /label= Transmembrane_domain
FT      141..154
FT      /label= Cytoplasmic_domain
FT      155..175
FT      /label= Transmembrane_domain
FT      176..181
FT      /label= Extracellular_domain
FT      /note= "Hydrophilic region useful for antibody
FT      production"
FT      182..198
FT      /label= Transmembrane_domain
FT      199..208
FT      /label= Cytoplasmic_domain
FT      209..232
FT      /label= Transmembrane_domain
FT      233..238
FT      /label= Extracellular_domain
FT      /note= "Hydrophilic region useful for antibody
FT      production"
FT      239..258
FT      /label= Transmembrane_domain
FT      259..314
FT      /label= Cytoplasmic_domain
FT      315..337
FT      /label= Transmembrane_domain
FT      338..375
FT      /label= Extracellular_domain
FT      /note= "Hydrophilic region useful for antibody
FT      production"
FT      376..394
FT      /label= Transmembrane_domain
FT      395..402
FT      /label= Cytoplasmic_domain
FT      403..421
FT      /label= Transmembrane_domain
FT      422..427

```

```

FT      /label= Extracellular_domain
FT      /note= "Hydrophilic region useful for antibody
FT      production"
FT      428..447
FT      /label= Transmembrane_domain
FT      448..457
FT      /label= Cytoplasmic_domain
FT      458..479
FT      /label= Transmembrane_domain
FT      480..486
FT      /label= Extracellular_domain
FT      /note= "Hydrophilic region useful for antibody
FT      production"
FT      487..510
FT      /label= Transmembrane_domain
FT      511..548
FT      /label= Cytoplasmic_domain
XX
PN      WO200000633-A1.
XX
XX      06-JAN-2000.
XX
XX      29-JUN-1999; 99WO-US014880.
XX
XX      30-JUN-1998; 98US-00107932.
XX      (MILL-) MILLENNIUM PHARM INC.
XX
XX      Goodearl AJ, Glucksmann MA;
XX      WPI; 2000-137069/12.
XX      N-PSDB; AAZ49686.
XX
XX      New nucleic acid encoding human organic cation transporter-like protein,
XX      used for prevention, treatment and diagnosis of e.g. neurological,
XX      behavioral or sleep disorders.
XX
XX      Claim 9; Fig 1; 100pp; English.
XX
XX      The present sequence is a human OCTlp (organic cation transporter-like
XX      protein), a member of the superfamily of sugar and other transporter
XX      molecules that have 12 transmembrane domains. The sequence is derived
XX      from a human foetal brain cDNA library. The protein is highly expressed
XX      in brain tissue and has nootropic, neuroprotective, neuroleptic,
XX      anticonvulsant, antiParkinsonian, antidepressant activities. The present
XX      sequence is used to regulate a variety of cellular processes e.g. cell
XX      proliferation, differentiation and survival, screen OCTlp modulators and
XX      prevent mutation in OCTlp gene. OCTlp modulators can be used to treat or
XX      prevent chronic neurodegenerative disorders (e.g. Alzheimer's,
XX      Parkinson's, Huntington's and ALS), CNS disorders (e.g. schizophrenia,
XX      panic, depression), behavioural, sleep and eating disorders
XX
XX      Sequence 548 AA;
XX
XX      Query Match      30.0%; Score 6; DB 3; Length 548;
XX      Best Local Similarity 100.0%; Pred. No. 2.8e+02;
XX      Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      8 LGWRWL 13
XX      |||||
XX      Db      235 LGWRWL 240
XX
XX      RESULT 39
XX      AAEL14907
XX      ID      AAEL14907 standard; protein; 548 AA.
XX      AC      AAEL14907;
XX      XX
XX      XX      27-AUG-2003 (first entry)
XX      DT
XX      DE      Human OCT-5 (57145) protein.
XX      XX

```

KW Pain disorder; inflammatory pain; chronic pain; neuropathic pain;
 KW causalgia; fibromyalgia; cancer pain; migraine; headache; tissue pain;
 KW cardiovascular disorder; analgesic; human; 57145; OCT-5.
 XX
 OS Homo sapiens.
 XX
 XX W02003039342-A2.
 XX
 XX 15-MAY-2003.
 XX
 XX 06-NOV-2002; 2002WO-US035562.
 XX
 XX 06-NOV-2001; 2001US-0333073P.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Silos-Santiago I;
 XX
 XX WPI; 2003-441438/41.
 DR N-PSDB; AAD36893.
 XX
 XX Treating a subject having a pain disorder characterized by aberrant 577,
 PT 20739 or 57145 polypeptide activity or nucleic acid expression, e.g.
 PT cancer pain or inflammatory pain, comprises administering a 577, 20739 or
 PT 57145 modulator.
 XX
 XX Claim 1; Page 83-84; 84pp; English.
 PS
 XX The present invention relates to a method of treating a subject having a
 CC pain disorder characterized by aberrant 577, 20739 or 57145 polypeptide
 CC activity or nucleic acid expression. The method involves administering a
 CC modulator of 577 (also known as sodium dependent proline transporter),
 CC 20739 (also known as P21-activated kinase 3 (PAK-3)) or 57145 (also known
 CC as OCT-5). The 577, 20739 or 57045 modulator is useful for treating pain
 CC disorders like inflammatory pain, chronic pain, neuropathic pain,
 CC causalgia, fibromyalgia, cancer pain, migraine/headache pain and tissue
 CC pain. The invention also describes methods for diagnosis and prognosis of
 CC various cardiovascular disorders, and for identification of subjects
 CC exhibiting a predisposition to such conditions. The present sequence is
 CC human 57145 protein
 XX
 XX Sequence 548 AA;
 SQ

Query Match 30.0%; Score 6; DB 6; Length 548;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 LGWRWL 13
 |||||
 Db 235 LGWRWL 240

RESULT 40
 ADO77835
 ID ADO77835 standard; protein; 548 AA.
 XX
 AC ADO77835;
 XX
 XX 26-AUG-2004 (first entry)
 DT
 XX
 XX Rat synaptic vesicle protein 2, SVOP.
 DE
 XX rat; levitiracetam; epilepsy; epileptogenesis; seizure disorder;
 KW convulsion; depression; anxiety; cerebral ischaemia; myotonia; stroke;
 KW Tourette's syndrome; neonatal cerebral haemorrhage; Parkinson's disease;
 KW Alzheimer's disease; dementia; synaptic vesicle protein 2; SVOP.
 XX
 OS Rattus norvegicus.
 XX
 XX US2004106147-A1.
 FN
 XX
 XX 03-JUN-2004.
 PD
 XX

PF 03-DEC-2002; 2002US-00308163.
 XX
 PR 03-DEC-2002; 2002US-00308163.
 XX
 XX (UNIO) UCB SA.
 XX
 XX Lynch B, Nocka K, Fuks B;
 PI
 XX WPI; 2004-467256/44.
 DR N-PSDB; ADO77834.
 DR
 XX Identifying binding partner for synaptic vesicle protein 2 (SV2) for
 PT treating epilepsy, Parkinson's disease, Alzheimer's disease, involves
 PT determining binding partner which modulates binding of levitiracetam or
 PT its analog to SV2 protein.
 XX
 XX Disclosure; SEQ ID NO 16; 63pp; English.
 PS
 XX The invention relates to a method of identifying a binding partner for
 CC synaptic vesicle protein 2 (SV2), involves incubating SV2 protein with
 CC levitiracetam and a potential binding partner and determining if the
 CC potential binding partner modulates the binding of levitiracetam to the
 CC SV2 protein or fragment, thus identifying a binding partner for the SV2
 CC protein. The method is useful for identifying a binding partner for
 CC synaptic vesicle protein 2 and the identified binding partner is useful
 CC for treating epilepsy, epileptogenesis, seizure disorders, convulsions,
 CC depression, anxiety, cerebral ischaemia, myotonia, stroke, Tourette's
 CC syndrome, neonatal cerebral haemorrhage, Parkinson's disease, Alzheimer's
 CC disease and dementia. The present sequence represents the amino acid
 CC sequence of the rat synaptic vesicle protein 2, SVOP.
 XX
 XX Sequence 548 AA;
 SQ

Query Match 30.0%; Score 6; DB 8; Length 548;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 LGWRWL 13
 |||||
 Db 235 LGWRWL 240

RESULT 41
 ADO77827
 ID ADO77827 standard; protein; 548 AA.
 XX
 AC ADO77827;
 XX
 XX 26-AUG-2004 (first entry)
 DT
 XX
 XX Human synaptic vesicle protein 2, SVOP, DNA.
 DE
 XX human; levitiracetam; epilepsy; epileptogenesis; seizure disorder;
 KW convulsion; depression; anxiety; cerebral ischaemia; myotonia; stroke;
 KW Tourette's syndrome; neonatal cerebral haemorrhage; Parkinson's disease;
 KW Alzheimer's disease; dementia; synaptic vesicle protein 2; SVOP.
 XX
 OS Homo sapiens.
 XX
 XX US2004106147-A1.
 FN
 XX
 XX 03-JUN-2004.
 PD
 XX 03-DEC-2002; 2002US-00308163.
 PF
 XX 03-DEC-2002; 2002US-00308163.
 XX
 XX (UNIO) UCB SA.
 XX
 XX Lynch B, Nocka K, Fuks B;
 PI
 XX WPI; 2004-467256/44.
 DR N-PSDB; ADO77826.
 DR